24675

U.S. DEPARTMENT OF COMM.
Patent and Trademark Office

SEARCH REQUEST FORM

Requestor's Name: GAMBEL / 1644 Serial Number: 08/819669

Date: 9/12/(X) Phone: 308 3997 Art Unit: 3004

Search Topic:
Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevent citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevent claim(s).

SEQ + SEQ INTERFERENCE SEARCH

SEQ ID NO: 26 x2

8

3a - 9/12 - 39-43

MP2 - 9/12 -

173
174
176
171
181
182

?
o

REQUEST IN

STAFF USE ONLY

Date completed: 9/13/00
Searcher:
Terminal time: 25
Elapsed time:
CPU time: x 10
Total time:
Number of Searches:
Number of Databases:

Search Site
___ STIC
✓ CM-1
___ Pre-S

Type of Search
✓ N.A. Sequence 1
✓ A.A. Sequence 2
___ Structure
___ Bibliographic

Vendors
✓ IG
___ STN
___ Dialog
___ APS
___ Geninfo
___ SDC
___ DARC/Questel
___ Other

MP2 rsm

PTO-1590 (9-90)

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: September 13, 2000, 01:18:33 ; Search time 2504.47 Seconds
(without alignments)
9990.683 Million cell updates/sec

Title: US-08-819-669E-8
Perfect score: 5674
Sequence: 1 CCCGGGGCACCACTGGCATC.....TAATGATCTTGGGTGGATCC 5674

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5247842 seqs, 2204914090 residues

Total number of hits satisfying chosen parameters: 10495684

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: em_est20:*
50: em_est21:*
51: em_est22:*
52: gb_est30:*
53: gb_est31:*
54: gb_est32:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: em_est35:*
83: em_est36:*
84: em_est37:*
85: gb_est48:*
86: gb_est49:*
87: gb_est50:*
88: gb_est51:*
89: gb_est52:*
90: gb_est53:*
91: gb_est54:*
92: gb_est55:*
93: gb_gss1:*
94: gb_gss2:*
95: gb_gss3:*
96: gb_gss4:*
97: em_gss1:*
98: em_gss2:*
99: em_gss3:*
100: em_gss4:*
101: gb_gss5:*
102: gb_gss6:*
103: em_gss5:*
104: em_gss6:*
105: gb_gss7:*
106: gb_gss8:*
107: gb_gss9:*
108: em_gss7:*
109: em_gss8:*
110: gb_gss11:*
111: gb_gss10:*
112: em_gss9:*
113: em_gss10:*
114: em_gss11:*
115: em_gss12:*
116: gb_gss12:*

117: gb_gss13:*
118: gb_gss14:*
119: gb_gss15:*
120: gb_gss16:*
121: gb_gss17:*
122: gb_gss18:*
123: gb_gss19:*
124: em_gss13:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	427.4	7.5	869	44	AI798898	AI798898 we94a11.x
c 2	413	7.3	414	63	AW103876	AW103876 xd47e12.x
c 3	392.6	6.9	728	44	AI805537	AI805537 tx86c10.x
c 4	375.4	6.6	670	72	AW438674	AW438674 xt02b04.x
5	349.6	6.2	533	71	AW383186	AW383186 PM3-HT034
6	331	5.8	578	47	AL044464	AL044464 DKFZp434H
c 7	328.2	5.8	599	47	AL044465	AL044465 DKFZp434H
8	286	5.0	464	79	AW673548	AW673548 ba55b09.y
9	285	5.0	504	47	AL045429	AL045429 DKFZp434F
c 10	275.2	4.9	493	45	AI863022	AI863022 wm45d10.x
c 11	272.6	4.8	511	69	AW245614	AW245614 2822978.3
c 12	270.6	4.8	481	33	AA995045	AA995045 ou53e02.s
c 13	270.6	4.8	495	69	AW194089	AW194089 xm12a03.x
c 14	260.2	4.6	521	70	AW250219	AW250219 2822505.3
c 15	257.8	4.5	459	36	AI222439	AI222439 qh04h04.x
c 16	251.2	4.4	451	44	AI805352	AI805352 tx85c10.x
c 17	247.6	4.4	430	79	AW628104	AW628104 hh98d02.x
c 18	246.4	4.3	415	36	AI200443	AI200443 qf93b07.x
19	243.4	4.3	519	69	AW245872	AW245872 2822978.5
20	225.4	4.0	534	70	AW247017	AW247017 2822505.5
c 21	222.8	3.9	394	32	AA857809	AA857809 oi03g03.s
c 22	220.2	3.9	457	80	AW731700	AW731700 ba55b09.x
c 23	218.6	3.9	566	39	AI394145	AI394145 tg60e07.x
c 24	215.4	3.8	780	40	AI479347	AI479347 tm27d08.x
c 25	212.4	3.7	386	70	AW273455	AW273455 xr39h01.x
c 26	212.2	3.7	638	70	AW248864	AW248864 2820966.3
c 27	206.8	3.6	514	79	AW664428	AW664428 hi11h03.x
c 28	205.2	3.6	399	69	AW193638	AW193638 xm18f07.x
c 29	201	3.5	488	44	AI830281	AI830281 wj95b10.x
c 30	193.8	3.4	378	63	AW103946	AW103946 xd62h08.x
c 31	193.2	3.4	467	35	AI142010	AI142010 oo21e08.x
c 32	188.8	3.3	459	93	AQ017378	AQ017378 CIT-HSP-2
33	188.4	3.3	597	70	AW249285	AW249285 2820966.5
34	181.6	3.2	338	88	R06041	R06041 ye89f01.r1
c 35	181.6	3.2	630	46	AI954607	AI954607 wq34a12.x
36	178.6	3.1	256	89	T29724	T29724 EST92093 Hu
37	176.4	3.1	457	119	AZ070771	AZ070771 RPCI-23-4
38	175	3.1	387	88	R06042	R06042 ye89f02.r1
c 39	166	2.9	466	30	AA704513	AA704513 zj23a12.s
40	164	2.9	293	89	T29746	T29746 EST92986 Hu
41	163.2	2.9	345	71	AW356066	AW356066 38053 MAR
c 42	159	2.8	324	32	AA905896	AA905896 oj88a05.s
c 43	157.2	2.8	467	34	AI032153	AI032153 os76e10.s
44	155.4	2.7	441	88	R23773	R23773 yh34b02.r1
45	150.6	2.7	523	116	AQ838224	AQ838224 HS_4729_A

ALIGNMENTS

RESULT 1
AI798898/c
LOCUS AI798898 869 bp mRNA EST 18-DEC-1999
DEFINITION we94a11.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:2348732 3' similar to SW:MAG2_HUMAN P43356
MELANOMA-ASSOCIATED ANTIGEN 2 ;, mRNA sequence.
ACCESSION AI798898
VERSION AI798898.1 GI:5364370
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 869)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Apr 7, 1998 this sequence version replaced gi:3036723.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 933 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 466.
FEATURES Location/Qualifiers
source 1. .869
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2348732"
/clone_lib="Soares_NFL_T_GBC_S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NbHL19W, testis NHT, and B-cell NCI_CGAP_GCB1) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 251 a 206 c 169 g 239 t 4 others
ORIGIN

Query Match 7.5%; Score 427.4; DB 44; Length 869;
Best Local Similarity 75.4%; Pred. No. 4.5e-96;
Matches 669; Conservative 0; Mismatches 185; Indels 33; Gaps 10;

```
Qy   4484 ATTGCAATGGAGGGCGGCCATGCTCCTGAGGAGGAAATCTGGGAGGAGCTGAGTGTGATG 4543
          || |||||| |||||||||  ||| |||||| || | |||||| || ||||     || ||
Db    869 ATCGCAATAGAGGGCGGCTGTGCCCCTGAGCAGCACATCTGGAAGNAGCTAGTAATGTTG 810

Qy   4544 GAGGTGTATGATGGGAGGGAGCACAGTGCCTATGGGGAGCCCAGGAAGCT-GCTCACCCA 4602
          |||||| ||| || | | |   ||||| || |   | ||||| ||||| |||||  ||
Db    809 NAGGTGT-TGAGGGAAAGAAAGTCAGTGTCTTCGCACATCCCAGAAAGCTAGCTCATGCA 751

Qy   4603 AGATTTGGTGCAGGAAAAGTACCTGGAGTACGGCAGGTGCCGGACAGTGATCCCGCACGC 4662
          ||| ||||||||    |  |||||| ||| |||||||    | ||         ||     |
Db    750 TGATCTGGTGCAGACTANCTACCTGAAGTTCGGCAGGGTGCCGAGCAGTGATCCTGCATC 691

Qy   4663 TATGAGTTCCTGTGGGGTCCAAGGGCCCTCGCTGAAACCAGCTATGTGAAAGTCCTTGAG 4722
          ||  | |||| | |||| |||||| ||  |||| | ||||||||||||||||  |  |
Db    690 CTCGATGTTCTGTAGAGTCCCAGGGCCATCATTGAACCAAGCTATGTGAAAGTCNTGCAC 631

Qy   4723 TATGTGATCAAGGTCAGTGCAAGAGTTCGCTTTTTCTTCCCATCCCTGCGTGAAGCAGCT 4782
          ||    | ||| || ||| |     || | ||| || |||| |||||| ||||   |||
Db    630 CATACACTAAAGATCGGTGAAGACCCTCACATTTCCTACCCACCCCTGCATGAACGGGCT 571

Qy   4783 TTGAGAGAGGAGGAAGAGGGAGTCTGAGCATGAGTTGCAGCCAAGGCCAGTGGGAGGGGG 4842
          ||||||||||  |||||| |||||| ||||  |||||||||| ||||||||||||||||
Db    570 TTGAGAGAGGGAGAAGAGTGAGTCTCAGCACATGTTGCAGCCAGGGCCAGTGGGAGGGGG 511
```

```
Qy   4843 ACTGGGCCAGTGCACCTTCCAGGGCCGCGTCCAGCAGCTTCCCCTGCCTCGTGTGACATG 4902
          ||||||||||||||||||||||||| | ||||  ||||||| ||||||||||||| |||
Db    510 TCTGGGCCAGTGCACCTTCCAGGGCCCCATCCATTAGCTTCCACTGCCTCGTGTGATATG 451

Qy   4903 AGGCCCATTCTT--CACTCTGAAGAGAGCGGTCAGTGTTCTCAGTAGTAGGTTTCTGTTC 4960
          |||||||||| |  | || |||||||||| |||||  |||| || |||  ||||||||||
Db    450 AGGCCCATTCCTGCCTCTTTGAAGAGAGCAGTCAGCATTCTTAGCAGTGAGTTTCTGTTC 391

Qy   4961 TATTGGGTGACTTGGAGATTTATCTTTGTTCTCTTTTGGAATTGTTCAAATGTTTTTTTT 5020
          | |||| |||||| |||||||||||||||||  || ||||||||||||||||| ||  |||
Db    390 TGTTGGATGACTTTGAGATTTATCTTTGTTTCCTGTTGGAATTGTTCAAATG-TTCCTTT 332

Qy   5021 TAAGGGATGGTTGAATGAACTTCAGCATCCAAGTTTATGAATGACAGCAGTCACAC--AG 5078
          |||   ||||||| |||||||||||||||||||||||||||||| ||| ||||||||  ||
Db    331 TAACAAATGGTTGGATGAACTTCAGCATCCAAGTTTATGAATGGCAGTAGTCACACATAG 272

Qy   5079 TTCTGTGTATATAGTTTAAGGGTAAGAGTCTTGTGTTTTATTCAGATTGGGAAATCCATT 5138
          | |||| |||||||||| |||||||||| ||| ||||||||||||||||||||||||||
Db    271 TGCTGTTTATATAGTTTAGGGGTAAGAGTCCTGTTTTTTATTCAGATTGGGAAATCCATT 212

Qy   5139 CTATTTTGTGAATTG--GGATAATAACAGCAGTGGAATAAGTACTTAGAAATGT----GA 5192
          | ||||||||| |||    |||||||||||||||||||| ||| ||  ||  |    |
Db    211 CCATTTTGTGAGTTGTCACATAATAACAGCAGTGGAATATGTATTTGCCTATATTGTGAA 152

Qy   5193 AAAATGAGCAGTAAAATAGATGAGATAAAGAACTAAAGAAATTAAGAGATAGTCAATTCT 5252
            ||| |||||||||||| |||| | || ||||        | || ||||||| ||||||
Db    151 CGAATTAGCAGTAAAATACATGATACAAGGAAC--------TCAAAAGATAGTTAATTCT 100

Qy   5253 TGCCTTATACCTCAGTCTATTCTGTAAAATTTTTAAAGATATATGCATACCTGGATTTCC 5312
          |||||||||||||||||||| |||||||   ||||| |||| || |         |
Db     99 TGCCTTATACCTCAGTCTATTATGTAAAA---TTAAAAATATGTGTA---------TGTT 52

Qy   5313 TTGGCTTCTTTGAGAATGTAAGAGAAATTAAATCTGAATAAAGAATT 5359
          || | |||| |||||||| || ||||||||||||   ||   ||||
Db     51 TTAGATTCTCTGAGAATGCAAAAGAAATTAAATCTAGGTAGGTAATT 5
```

RESULT 2
AW103876/c
LOCUS AW103876 414 bp mRNA EST 20-OCT-1999
DEFINITION xd47e12.x1 NCI_CGAP_Ov23 Homo sapiens cDNA clone IMAGE:2596942 3', mRNA sequence.
ACCESSION AW103876
VERSION AW103876.1 GI:6074611
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 414)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco
High quality sequence stop: 399.
FEATURES Location/Qualifiers
source 1. .414
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2596942"
/clone_lib="NCI_CGAP_Ov23"
/tissue_type="tumor, 5 pooled (see description)"
/lab_host="DH10B"
/note="Organ: ovary; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.35 kb. Tumor types include: mixed Mullerian tumor, papillary serous, clear cell, spindle cell. All are primary tumors, metastasis positive. Life Technologies catalog #: 11534-013"
BASE COUNT 149 a 84 c 46 g 134 t 1 others
ORIGIN

Query Match 7.3%; Score 413; DB 63; Length 414;
Best Local Similarity 99.8%; Pred. No. 1.5e-92;
Matches 413; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy   4942 CAGTAGTAGGTTTCTGTTCTATTGGGTGACTTGGAGATTTATCTTTGTTCTCTTTTGGAA 5001
          |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    414 CAGTAGTAGGTTTCTGTTCTATTGGGTGACTTGNAGATTTATCTTTGTTCTCTTTTGGAA 355

Qy   5002 TTGTTCAAATGTTTTTTTTTAAGGGATGGTTGAATGAACTTCAGCATCCAAGTTTATGAA 5061
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    354 TTGTTCAAATGTTTTTTTTAAGGGATGGTTGAATGAACTTCAGCATCCAAGTTTATGAA 295

Qy   5062 TGACAGCAGTCACACAGTTCTGTGTATATAGTTTAAGGGTAAGAGTCTTGTGTTTTATTC 5121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    294 TGACAGCAGTCACACAGTTCTGTGTATATAGTTTAAGGGTAAGAGTCTTGTGTTTTATTC 235

Qy   5122 AGATTGGGAAATCCATTCTATTTTGTGAATTGGGATAATAACAGCAGTGGAATAAGTACT 5181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    234 AGATTGGGAAATCCATTCTATTTTGTGAATTGGGATAATAACAGCAGTGGAATAAGTACT 175

Qy   5182 TAGAAATGTGAAAAATGAGCAGTAAAATAGATGAGATAAAGAACTAAAGAAATTAAGAGA 5241
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    174 TAGAAATGTGAAAAATGAGCAGTAAAATAGATGAGATAAAGAACTAAAGAAATTAAGAGA 115

Qy   5242 TAGTCAATTCTTGCCTTATACCTCAGTCTATTCTGTAAAATTTTTAAAGATATATGCATA 5301
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    114 TAGTCAATTCTTGCCTTATACCTCAGTCTATTCTGTAAAATTTTTAAAGATATATGCATA 55

Qy   5302 CCTGGATTTCCTTGGCTTCTTTGAGAATGTAAGAGAAATTAAATCTGAATAAAG 5355
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     54 CCTGGATTTCCTTGGCTTCTTTGAGAATGTAAGAGAAATTAAATCTGAATAAAG 1
```

RESULT 3
AI805537/c
LOCUS AI805537 728 bp mRNA EST 16-DEC-1999
DEFINITION tx86c10.x1 NCI_CGAP_Ut4 Homo sapiens cDNA clone IMAGE:2276466 3' similar to SW:MAG4_HUMAN P43358 MELANOMA-ASSOCIATED ANTIGEN 4 ;, mRNA sequence.
ACCESSION AI805537
VERSION AI805537.1 GI:5392103
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 728)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On May 11, 1999 this sequence version replaced gi:4776345.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be

```
          found through the I.M.A.G.E. Consortium/LLNL at:
          www-bio.llnl.gov/bbrp/image/image.html
          Insert Length: 1984   Std Error: 0.00
          Seq primer: -40UP from Gibco
          High quality sequence stop: 404.
FEATURES             Location/Qualifiers
     source          1. .728
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2276466"
                     /clone_lib="NCI_CGAP_Ut4"
                     /tissue_type="serous papillary carcinoma, high grade, 2
                     pooled tumors"
                     /lab_host="DH10B"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.48 kb. Life Technologies catalog #:
                     11542-016"
BASE COUNT      234 a    179 c    119 g    192 t      4 others
ORIGIN
```

```
  Query Match             6.9%;  Score 392.6;  DB 44;  Length 728;
  Best Local Similarity   80.8%;  Pred. No. 2.2e-87;
  Matches  600;  Conservative    0;  Mismatches  108;  Indels   35;  Gaps   11;

Qy    4638 GGTGCCGGACAGTGATCCCGCACGCTATGAGTTCCTGTGGGGTCCAAGGGCCCTCGCTGA 4697
           ||| ||||  ||| |||| || |||||||||||| ||||||||||||||| || ||||
Db     720 GGTACCGGCAAGTAATCCTGCGCGCTATGAGTTCTTGTGGGGTCCAAGGGCTCTGGCTG- 660

Qy    4698 AACCAGCTATGTGAAAGTCCTTGAGTATGTGATCAAGGTCAGTGCAAGAGTTCGCTTTTT 4757
           ||||||  ||||||||||||    || ||||| ||| ||||| ||||||||||||  ||
Db    -661 AACCAGTAATGTGAAAGTCTGNNAGCATGTGGTCAGGGTCAATGCAAGAGTTCGC-ATTG 603

Qy    4758 CTTCCCATCCCTGCGTGAAGCAGCTTTGAGAGAGGAGGAAGAGGGAGTCTGAGCATGAGT 4817
           || |||||||   |||||||||||| || ||||||||||||||||||||||||||||||
Db     602 CTACCCATCCTGNCGTGAAGCAGCTNTGTTAGAGGAGGAAGAGGGAGTCTGAGCATGAGT 543

Qy    4818 TGCAGCCAAGGCCAGTGGGAGGG----GGACTGGGCCAGTGCACCTTCCAGGGCCGCGTC 4873
             ||| |||    |||| ||    || |||||||||||| ||  ||  |||  ||
Db     542 GCAGCCCAGGGCTGTGGGGAAGGGGCAGGGCTGGGCCAGTGCATCTAACA--GCCCTGTG 485

Qy    4874 CAGCAGCTTCCCCTGCCTCGTGTGACATGAGGCCCATTCTTCACTC----TGAAGAGAGC 4929
           |||||||||||| |||||||||| ||||||||||||||||||||    |||||| |
Db     484 CAGCAGCTTCCCTTGCCTCGTGTAACATGAGGCCCATTCTTCACTCTGTTTGAAGAAAAT 425

Qy    4930 GGTCAGTGTTCTCAGTAGTAGGTTTCTGTTCTATTGGGTGACTTGGAGATTTATCTTTGT 4989
           |||||||||||| |||||| ||||||| || | |||| ||||||||||||||||||| |||
Db     424 AGTCAGTGTTCTTAGTAGTGGGTTTCTATTTTGTTGGATGACTTGGAGATTTATCTCTGT 365

Qy    4990 TCTCTTTTGGAATTGTTCAAATGTTTTTTTTTAAGGGATGGTTGAATGAACTTCAGCATC 5049
           |  |||||  |||||||  |||||||    ||||||  |||||||||||| ||||||||||
Db     364 TTCCTTTTACAATTGTTGAAATGTT-CCTTTTAATGGATGGTTGAATTAACTTCAGCATC 306

Qy    5050 CAAGTTTATGAATGACAGCAGTCACACAGTTCTGTGTATATAGTTTAAGGGTAAGAGTCT 5109
           |||||||||||||   ||    |  | | | ||||  |||||||||| | ||||||||||
Db     305 CAAGTTTATGAATCGTAGTTAACGTATATTGCTGTTAATATAGTTTAGGAGTAAGAGTCT 246

Qy    5110 TGTGTTTTATTCAGATTGGGAAATCCATTCTATTTTGTGAATTGGG---ATAATAACAGC 5166
           ||| |||||||||||||||||||||| ||||||||||||||||  ||   ||||||||||
Db     245 TGTTTTTTATTCAGATTGGGAAATCCGTTCTATTTTGTGAATTTGGGACATAATAACAGC 186

Qy    5167 AGTGGAATAAGTACTTAGAAATGTGAAAAATGAGCAGTAAAATAGATGAGATAAAGAACT 5226
           ||||||  ||||||  ||||    |||    | || ||||||  ||||||
Db     185 AGTGGAGTAAGTATTTAGAAGTGTG---AATTCACCGTGAAATAGGTGAGAT-------- 135

Qy    5227 AAAGAAATTAAGAGATAGTCAATTCTTGCCTTATACCTCAGTCTATTCTGTAAAATTT-T 5285
           |||||||  |||||  | |||||  |||||||  ||||||||||||||||||||||| |
Db     136 ----AAATTAAAAGATACTTAATTCCCGCCTTATGCCTCAGTCTATTCTGTAAAATTTAA 81

Qy    5286 TAAAGATATATGCATACCTGGATTTCCTTGGCTTCTTTGAGAATGTAAGAGAAATTAAAT 5345
           ||| |||||||||||||||||||||||||||||    | ||||||||||||||||||||
Db      80 AAAATATATATGCATACCTGGATTTCCTTGGCTTC---GTGAATGTAAGAGAAATTAAAT 24
```

```
Qy    5346 CTGAATAAAGAATTCTTCCTGTT 5368
           ||||||||| ||||||| |||||
Db      23 CTGAATAAATAATTCTTTCTGTT 1


RESULT   4
AW438674/c
LOCUS       AW438674      670 bp    mRNA            EST       14-FEB-2000
DEFINITION  xt02b04.x1 NCI_CGAP_Ut4 Homo sapiens cDNA clone IMAGE:2777935 3'
            similar to SW:MAG4_HUMAN P43358 MELANOMA-ASSOCIATED ANTIGEN 4 ;,
            mRNA sequence.
ACCESSION   AW438674
VERSION     AW438674.1  GI:6973980
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 670)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     On Jan 6, 2000 this sequence version replaced gi:6677137.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40UP from Gibco
            High quality sequence stop: 420.
FEATURES             Location/Qualifiers
     source          1. .670
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2777935"
                     /clone_lib="NCI_CGAP_Ut4"
                     /tissue_type="serous papillary carcinoma, high grade, 2
                     pooled tumors"
                     /lab_host="DH10B"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.48 kb. Life Technologies catalog #:
                     11542-016"
BASE COUNT      219 a    155 c    109 g    187 t
ORIGIN
```

```
  Query Match             6.6%;  Score 375.4;  DB 72;  Length 670;
  Best Local Similarity   81.8%;  Pred. No. 4.4e-83;
  Matches  552;  Conservative    0;  Mismatches   91;  Indels   32;  Gaps    9;

Qy    4707 TGTGAAAGTCCTTGAGTATGTGATCAAGGTCAGTGCAAGAGTTCGCTTTTTCTTCCCATC 4766
           |||||||||||| ||| ||||| ||| ||||| ||||||||| ||| ||  || ||||||
Db     659 TGTGAAAGTCCTGGAGCATGTGGTCAGGGTCAATGCAAGAGTCCGCATTGCCTACCCATC 600

Qy    4767 CCTGCGTGAAGCAGCTTTGAGAGAGGAGGAAGAGGGAGTCTGAGCATGAGTTGCAGCCAA 4826
           |||||||||| ||||||||  |||| || ||||||||||||||||||||| |||||||
Db     599 CCTGCGTGAACCAGCTTTGTTAGAGAAGCAAGAGGGAGTCTGAGCATGAGTGGCAGCCAG 540

Qy    4827 GGCCAGTGGGAGGG---GGACTGGGCCAGTGCACCTTCCAGGGCCGCGTCCAGCAGCTTC 4883
           |||    |||||||   || |||||||||||| ||  ||  |||  || ||||||||||
Db     539 GGCTGTGGGGAGGGGCAGGGCTGGGCCAGTGCATCTAACA--GCCCTGTGCAGCAGCTTC 482

Qy    4884 CCCTGCCTCGTGTGACATGAGGCCCATTCTTCACTC----TGAAGAGAGCGGTCAGTGTT 4939
           || |||||||||| |||||||||||||||||||||    |||||| |  |||||||||
```

```
Db    481 CCTTGCCTCGTGTAACATGAGGCCCATTCTTCACTCTGTTTGAAGAAAATAGTCAGTGTT 422

Qy   4940 CTCAGTAGTAGGTTTCTGTTCTATTGGGTGACTTGGAGATTTATCTTTGTTCTCTTTTGG 4999
          || |||||| ||||||| || | |||| |||||||||||||||||| ||||  |||||
Db    421 CTTAGTAGTGGGTTTCTATTTTGTTGGATGACTTGGAGATTTATCTCTGTTTCCTTTTAC 362

Qy   5000 AATTGTTCAAATGTTTTTTTTTAAGGGATGGTTGAATGAACTTCAGCATCCAAGTTTATG 5059
          ||||||| ||||| ||  |||||| |||||||||||| ||||||||||||||||||||| |
Db    361 AATTGTTGAAATG-TTCCTTTTAATGGATGGTTGAATTAACTTCAGCATCCAAGTTTATG 303

Qy   5060 AATGACAGCAGTCACACAGTTCTGTGTATATAGTTTAAGGGTAAGAGTCTTGTGTTTTAT 5119
          |||   ||    | | | | |||| |||||||||| | ||||||||||||| ||||||
Db    302 AATCGTAGTTAACGTATATTGCTGTTAATATAGTTTAGGAGTAAGAGTCTTGTTTTTTAT 243

Qy   5120 TCAGATTGGGAAATCCATTCTATTTTGTGAATTGGG---ATAATAACAGCAGTGGAATAA 5176
          |||||||||||||||| |||||||||||||||| ||   |||||||||||||||| |||
Db    242 TCAGATTGGGAAATCCGTTCTATTTTGTGAATTTGGGACATAATAACAGCAGTGGAGTAA 183

Qy   5177 GTACTTAGAAATGTGAAAAATGAGCAGTAAAATAGATGAGATAAAGAACTAAAGAAATTA 5236
          ||| |||||| |||| |||  | || |||||| ||||||             ||||||
Db    182 GTATTTAGAAGTGTG---AATTCACCGTGAAATAGGTGAGAT-----------AAATTA 138

Qy   5237 AGAGATAGTCAATTCTTGCCTTATACCTCAGTCTATTCTGTAAAATTT-TTAAAGATATA 5295
          | ||||| | ||||| |||||||| ||||||||||||||||||||||   ||| |||||
Db    137 AAAGATACTTAATTCCCGCCTTATGCCTCAGTCTATTCTGTAAAATTTAAAAAATATATA 78

Qy   5296 TGCATACCTGGATTTCCTTGGCTTCTTTGAGAATGTAAGAGAAATTAAATCTGAATAAAG 5355
          |||||||||||||||||||||||    | ||||||||||||||||||||||||||||||
Db     77 TGCATACCTGGATTTCCTTGGCTTC---GTGAATGTAAGAGAAATTAAATCTGAATAAAT 21

Qy   5356 AATTCTTCCTGTTCA 5370
          ||||||| ||||| |
Db     20 AATTCTTTCTGTTAA 6


RESULT   5
AW383186
LOCUS       AW383186      533 bp    mRNA            EST       04-FEB-2000
DEFINITION  PM3-HT0344-151299-004-h01 HT0344 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW383186
VERSION     AW383186.1  GI:6887845
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 533)
  AUTHORS   HCGP http://www.ludwig.org.br/ORESTES.
  TITLE     The FAPESP/LICR Human Cancer Genome Project
  JOURNAL   Unpublished (1999)
COMMENT     On Jun 15, 1998 this sequence version replaced gi:3224202.
            Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=PM3&t2=PM3-HT0344-
            151299-004-h01&t3=1999-12-15&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 21
            High quality sequence stop: 533.
FEATURES             Location/Qualifiers
     source          1. .533
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="HT0344"
                     /dev_stage="Adult"
                     /note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT      117 a    112 c    175 g    129 t
ORIGIN


  Query Match             6.2%;  Score 349.6;  DB 71;  Length 533;
  Best Local Similarity   84.0%;  Pred. No. 1.2e-76;
  Matches  445;  Conservative    0;  Mismatches   74;  Indels   11;  Gaps    4;

Qy   4484 ATTGCAATGGAGGGCGGCCATGCTCCTGAGGAGGAAATCTGGGAGGAGCTGAGTGTGATG 4543
          ||| |  |  ||  | |   ||  |||||||||||||||||||||||||||||| ||||||||
Db      6 ATTCAATGGAGGGCTGACAGCGCCTCTGAGGAGGAAATCTGGGAGGAGCTGGGTGTGATG 65

Qy   4544 GAGGTGTATGATGGGAGGGAGCACAGTGCCTATGGGGAGCCCAGGAAGCTGCTCACCCAA 4603
          | |||||||||||||||||||||||| || |||||||||||||||||| ||||||||||||
Db     66 GGGGTGTATGATGGGAGGGAGCACACTGTCTATGGGGAGCCCAGGAAACTGCTCACCCAA 125

Qy   4604 GATTTGGTGCAGGAAAAGTACCTGGAGTA-CGGCAGGTGCCGGACAGTGATCCCGCACGC 4662
          |||| |||||||||||| |||||||||| |||||||| || | |||| |||| || |||
Db    126 GATTGGGTGCAGGAAAACTACCTGGAGTACCGGCAGGTACCCGGCAGTAATCCTGCGCGC 185

Qy   4663 TATGAGTTCCTGTGGGGTCCAAGGGCCCTCGCTGAAACCAGCTATGTGAAAGTCCTTGAG 4722
          ||||||||||||||||||||||||||| || |||||||||||||||||||||||||| |||
Db    186 TATGAGTTCCTGTGGGGTCCAAGGGCTCTGGCTGAAACCAGCTATGTGAAAGTCCTGGAG 245

Qy   4723 TATGTGATCAAGGTCAGTGCAAGAGTTCGCTTTTTCTTCCCATCCCTGCGTGAAGCAGCT 4782
           ||||| ||| |||||| |||||||||||| ||  || ||||||||||||||||||||||
Db    246 CATGTGGTCAGGGTCAATGCAAGAGTTCGCATTGCCTACCCATCCCTGCGTGAAGCAGCT 305

Qy   4783 TTGAGAGAGGAGGAAGAGGGAGTCTGAGCATGAGTTGCAGCCAAGGCCAGTGGGAGGG-- 4840
          |||  |||||||||||||||||||||||||||||||||||||||||||| |||    |||| ||
Db    306 TTGTTAGAGGAGGAAGAGGGAGTCTGAGCATGAGTTGCAGCCAGGGCTGTGGGGAAGGGG 365

Qy   4841 --GGACTGGGCCAGTGCACCTTCCAGGGCCGCGTCCAGCAGCTTCCCCTGCCTCGTGTGA 4898
            || |||||||||||||| ||  ||  |||  || ||||||||||| |||||||||| |
Db    366 CAGGGCTGGGCCAGTGCATCTAACA--GCCCTGTGCAGCAGCTTCCCTTGCCTCGTGTAA 423

Qy   4899 CATGAGGCCCATTCTTCACTC----TGAAGAGAGCGGTCAGTGTTCTCAGTAGTAGGTTT 4954
          ||||||||||||||||||||    ||| || |  |||||||||| |||||| ||||
Db    424 CATGAGGCCCATTCTTCACTCTGTTTGAGGAAAATAGTCAGTGTTCTTAGTAGTGGGTTA 483

Qy   4955 CTGTTCTATTGGGTGACTTGGAGATTTATCTTTGTTCTCTTTTGGAATTG 5004
          || ||   |||| |||||||||||||||| ||||  | |||  |||||
Db    484 CTATTTGGTTGGATGACTTGGAGATTTATCTCTGTTTCCCTTTACAATTG 533


RESULT   6
AL044464
LOCUS       AL044464      578 bp    mRNA            EST       29-FEB-2000
DEFINITION  DKFZp434H022_r1 434 (synonym: htes3) Homo sapiens cDNA clone
            DKFZp434H022 5', mRNA sequence.
ACCESSION   AL044464
VERSION     AL044464.1  GI:5432682
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 578)
  AUTHORS   Ansorge,W., Benes,V., Krieger,S., Mewes,H.W., Gassenhuber,J. and
            Wiemann,S.
  TITLE     EST (Ansorge, Benes, et al.)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Ansorge W
            MIPS
            Am Klopferspitz 18a D-82152 Martinsried, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
```

```
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by EMBL (European Molecular Biology Laboratories,
            Heidelberg/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            s1 sequence also available.
            This clone (DKFZp434H022) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES              Location/Qualifiers
     source           1. .578
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="DKFZp434H022"
                      /clone_lib="434 (synonym: htes3)"
                      /tissue_type="testis"
                      /dev_stage="adult"
                      /lab_host="DH10B"
                      /note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT      130 a    136 c    196 g    116 t
ORIGIN


  Query Match             5.8%;  Score 331;  DB 47;  Length 578;
  Best Local Similarity   78.0%;  Pred. No. 5.4e-72;
  Matches  453;  Conservative    0;  Mismatches  115;  Indels   13;  Gaps    4;

Qy   2531 GTTGGGGGGCCCTCAGGGAGATGGGGTCTTGGGGTAAAGGGGGGATGTCTACTCATGTCA 2590
           | | ||||  ||||||||||||  ||| |||| |||||| || | ||||| |||||   |
Db      1 GGTTGGGGAACCTCAGGGAGATAAGGTGTTGGTGTAAAGAGGAGCTGTCTGCTCATT--A 58

Qy   2591 GGGAATTGGGGGTTGAGGAAGCACAGGCGCTGGCAGGAATAAAGATGAGTGAGACAGACA 2650
           |||  ||||||||||| |||  ||| | |||||||| ||||||||||| |  || |
Db     59 GGGGGTTGGGGGTTGAGAAAGGGCAGTCCCTGGCAGGAGTAAAGATGAGTAACCCACAGG 118

Qy   2651 AGGCTATTGGAATCCACACCCCAGAACCAAAGGGGTCAGCCCTGGACACCTCAC------ 2704
           |||| ||   ||    ||||| |||||||||||||||||||||||||| | |||
Db    119 AGGCCATCATAACGTTCACCCTAGAACCAAAGGGGTCAGCCCTGGACAACGCACGTGGGG 178

Qy   2705 ---CCAGGATGTGGCTTCTTTTTCACTCCTGTTTCCAGATCTGGGGCAGGTGAGGACCTC 2761
              |||||||||| | | ||||| |||||||||||||  || || ||| |||||
Db    179 GTAACAGGATGTGGC-CCCTCCTCACTTCTGTTTCCAGATCTCAGGGAGTTGATGACCTT 237

Qy   2762 ATTCTCAGAGGGTGACTCAGGTCAACGTAGGGACCCCCATCTGGTCTAAAGACAGAGCGG 2821
           || ||||| |||||||||||||||  |||| |||||||||||| | |||   || ||
Db    238 GTTTTCAGAAGGTGACTCAGGTCAACACAGGGGCCCCCATCTGGTCGACAGATGCAGTGG 297

Qy   2822 TCCCAGGATCTGCCATGCGTTCGGGTGAGGAACATGAGGGAGGACTGAGGGTACCCCAGG 2881
           | | |||||||||||| || | | |||| || | ||||| |||| ||||||||||||
Db    298 TTCTAGGATCTGCCAAGCATCCAGGTGGAGAGCCTGAGGTAGGATTGAGGGTACCCCTGG 357

Qy   2882 ACCAGAACACTGA-GGGAGACTGCACAGAAATCAGCCCTGCCCCTGCTGTCACCCCAGAG 2940
           ||||||  | ||  || | |  || ||||||| |||||||||||| || ||  |||||
Db    358 GCCAGAATGCAGACAGGGGGCCCCATAGAAATCTGCCCTGCCCCTGCGGTTACTTCAGAG 417

Qy   2941 AGCATGGGCTGGGCCGTCTGCCGAGGTCCTTCCGTTATCCTGGGATCATTGATGTCAGGG 3000
           | | ||||| |||| ||| || || |||| ||| |||||||||||| ||||||||||
Db    418 ACCCTGGGCAGGGCTGTCAGCTGAAGTCCCTCCATTATCCTGGGATCTTTGATGTCAGGG 477

Qy   3001 ACGGGGAGGCCTTGGTCTGAGAAGGCTGCGCTCAGGTCAGTAGAGGGAGCGTCCCAGGCC 3060
           | |||||||||||||||||    |||||   |||||||||||||||||| ||| ||||||
Db    478 AAGGGGAGGCCTTGGTCTGAAGGGGCTGGAGTCAGGTCAGTAGAGGGAGGGTCTCAGGCC 537

Qy   3061 CTGCCAGGAGTCAAGGTGAGGACCAAGCGGGCACCTCACCC 3101
           ||||||||||  | |||||||||||||| | | ||||||
Db    538 CTGCCAGGAGTGGACGTGAGGACCAAGCGGACTCGTCACCC 578


RESULT   7
AL044465/c
LOCUS       AL044465      599 bp    mRNA            EST       29-FEB-2000
DEFINITION  DKFZp434H022_s1 434 (synonym: htes3) Homo sapiens cDNA clone
            DKFZp434H022 3', mRNA sequence.
ACCESSION   AL044465
VERSION     AL044465.1  GI:5432683
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 599)
  AUTHORS   Ansorge,W., Benes,V., Krieger,S., Mewes,H.W., Gassenhuber,J. and
            Wiemann,S.
  TITLE     EST (Ansorge, Benes, et al.)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Ansorge W
            MIPS
            Am Klopferspitz 18a D-82152 Martinsried, Germany
            This is the 3' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by EMBL (European Molecular Biology Laboratories,
            Heidelberg/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            r1 sequence also available.
            This clone (DKFZp434H022) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES              Location/Qualifiers
     source           1. .599
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="DKFZp434H022"
                      /clone_lib="434 (synonym: htes3)"
                      /tissue_type="testis"
                      /dev_stage="adult"
                      /lab_host="DH10B"
                      /note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT      191 a    130 c     98 g    180 t
ORIGIN


  Query Match             5.8%;  Score 328.2;  DB 47;  Length 599;
  Best Local Similarity   81.1%;  Pred. No. 2.8e-71;
  Matches  489;  Conservative    0;  Mismatches   83;  Indels   31;  Gaps    8;

Qy   4766 CCCTGCGTGAAGCAGCTTTGAGAGAGGAGGAAGAGGGAGTCTGAGCATGAGTTGCAGCCA 4825
           ||||||  ||||   |||||||||||||  |||||| |||||| ||||   ||||||||||
Db    599 CCCTGCATGAACGGGCTTTGAGAGAGGGAGAAGAGTGAGTCTCAGCACATGTTGCAGCCA 540

Qy   4826 AGGCCAGTGGGAGGGGGACTGGGCCAGTGCACCTTCCAGGGCCGCGTCCAGCAGCTTCCC 4885
           ||||||||||||||||  |||||||||||||||||||||||||| | ||||  |||||||
Db    539 GGGCCAGTGGGAGGGGGTCTGGGCCAGTGCACCTTCCAGGGCCCCATCCATTAGCTTCCA 480

Qy   4886 CTGCCTCGTGTGACATGAGGCCCATTCTT--CACTCTGAAGAGAGCGGTCAGTGTTCTCA 4943
           ||||||||||||| |||||||||||||  | || ||||||||||  |||||  |||| |
Db    479 CTGCCTCGTGTGATATGAGGCCCATTCCTGCCTCTTTGAAGAGAGCAGTCAGCATTCTTA 420

Qy   4944 GTAGTAGGTTTCTGTTCTATTGGGTGACTTGGAGATTTATCTTTGTTCTCTTTTGGAATT 5003
           | |||  |||||||||||  |||| ||||||  |||||||||||||||||    | ||||||||
Db    419 GCAGTGAGTTTCTGTTCTGTTGGATGACTTTGAGATTTATCTTTGTTTCATGTTGGAATT 360

Qy   5004 GTTCAAATGTTTTTTTTTTAAGGGATGGTTGAATGAACTTCAGCATCCAAGTTTATGAATG 5063
           |||||||| ||  ||||||   ||||||| ||||||||||||||||||||||||||||||
Db    359 GTTCAAATG-TTCCTTTTAACAAATGGTTGGATGAACTTCAGCATCCAAGTTTATGAATG 301

Qy   5064 ACAGCAGTCACAC--AGTTCTGTGTATATAGTTTAAGGGTAAGAGTCTTGTGTTTTATTC 5121
           |||| ||||||||  ||| |||| |||||||||| |||||||||| ||| ||||||||
Db    300 ACAGTAGTCACACATAGTGCTGTTTATATAGTTTAGGGGTAAGAGTCCTGTTTTTTATTC 241

Qy   5122 AGATTGGGAAATCCATTCTATTTTGTGAATTG--GGATAATAACAGCAGTGGAATAAGTA 5179
           |||||||||||||||||| |||||||| |||    |||||||||||||||||||| |||
Db    240 AGATTGGGAAATCCATTCCATTTTGTGAGTTGTCACATAATAACAGCAGTGGAATATGTA 181

Qy   5180 CTTAGAAATGT----GAAAAATGAGCAGTAAAATAGATGAGATAAAGAACTAAAGAAATT 5235
           ||    || |    | ||| |||||||||||| |||| | || ||||       |
Db    180 TTTGCCTATATTGTGAACGAATTAGCAGTAAAATACATGATACAAGGAAC--------TC 129
```

```
Qy   5236 AAGAGATAGTCAATTCTTGCCTTATACCTCAGTCTATTCTGTAAAATTTTTAAAGATATA 5295
          || ||||||| |||||||||||||||||||||||||||||| |||||||   ||||| ||||
Db    128 AAAAGATAGTTAATTCTTGCCTTATACCTCAGTCTATTATGTAAAA---TTAAAAATATG 72

Qy   5296 TGCATACCTGGATTTCCTTGGCTTCTTTGAGAATGTAAGAGAAATTAAATCTGAATAAAG 5355
          || |         |   || |||||||||||||| || ||||||||||||||||||||
Db     71 TGTA---------TGTTTTTGCTTCTTTGAGAATGCAAAAGAAATTAAATCTGAATAATA 21

Qy   5356 AAT 5358
          | |
Db     20 ATT 18


RESULT   8
AW673548
LOCUS       AW673548      464 bp    mRNA            EST       11-APR-2000
DEFINITION  ba55b09.y3 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:2900441 5',
            mRNA sequence.
ACCESSION   AW673548
VERSION     AW673548.1  GI:7538783
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 464)
  AUTHORS   NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     On Jul 7, 1999 this sequence version replaced gi:5406105.
            Other_ESTs: ba55b09.x1
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            image.llnl.gov/image/html/iresources.shtml
            Seq primer: -40RP from Gibco
            High quality sequence stop: 439.
FEATURES             Location/Qualifiers
     source          1. .464
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2900441"
                     /clone_lib="NIH_MGC_10"
                     /cell_line="MGC36"
                     /lab_host="DH10B"
                     /note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.5 kb. Library prepared by Life
                     Technologies."
BASE COUNT      122 a     82 c    113 g    147 t
ORIGIN


  Query Match             5.0%;  Score 286;  DB 79;  Length 464;
  Best Local Similarity   83.5%;  Pred. No. 9.4e-61;
  Matches  386;  Conservative    0;  Mismatches   65;  Indels   11;  Gaps    5;

Qy   4783 TTGAGAGAGGAGGAAGAGGGAGTCTGAGCATGAGTTGCAGCCAAGGCCAGTGGGAGGGGG 4842
          ||||||||||  |||||| |||||| ||||   |||||||||| ||||||||||||||||
Db      1 TTGAGAGAGGGAGAAGAGTGAGTCTCAGCACATGTTGCAGCCAGGGCCAGTGGGAGGGGG 60

Qy   4843 ACTGGGCCAGTGCACCTTCCAGGGCCGCGTCCAGCAGCTTCCCCTGCCTCGTGTGACATG 4902
           |||||||||||||||||||||||||| | ||||  ||||||| |||||||||||||| |||
Db     61 TCTGGGCCAGTGCACCTTCCAGGGCCCCATCCATTAGCTTCCACTGCCTCGTGTGATATG 120

Qy   4903 AGGCCCATTCTT--CACTCTGAAGAGAGCGGTCAGTGTTCTCAGTAGTAGGTTTCTGTTC 4960
          |||||||||| |  | || |||||||||| |||||  |||| || |||  ||||||||||
Db    121 AGGCCCATTCCTGCCTCTTTGAAGAGAGCAGTCAGCATTCTTAGCAGTGAGTTTCTGTTC 180

Qy   4961 TATTGGGTGACTTGGAGATTTATCTTTGTTCTCTTTTGGAATTGTTCAAATGTTTTTTTT 5020
          | |||| |||||| |||||||||||||||  || ||||||||||||||||| ||  |||
Db    181 TGTTGGATGACTTTGAGATTTATCTTTGTTTCCTGTTGGAATTGTTCAAATG-TTCCTTT 239

Qy   5021 TAAGGGATGGTTGAATGAACTTCAGCATCCAAGTTTATGAATGACAGCAGTCACAC--AG 5078
          |||   ||||||| |||||||||||||||||||||||||||||||| ||||||||  ||
Db    240 TAACAAATGGTTGGATGAACTTCAGCATCCAAGTTTATGAATGACAGTAGTCACACATAG 299

Qy   5079 TTCTGTGTATATAGTTTAAGGGTAAGAGTCTTGTGTTTTATTCAGATTGGGAAATCCATT 5138
          | |||| |||||||||||| |||||||||| ||| ||||||||||||||||||||||||
Db    300 TGCTGTTTATATAGTTTAGGGGTAAGAGTCCTGTTTTTTATTCAGATTGGGAAATCCATT 359

Qy   5139 CTATTTTGTGAATTG--GGATAATAACAGCAGTGGAATAAGTACTTAGAAATGT----GA 5192
          | ||||||||| |||    |||||||||||||||||||| ||| ||    || |    |
Db    360 CCATTTTGTGAGTTGTCACATAATAACAGCAGTGGAATATGTATTTGCCTATATTGTGAA 419

Qy   5193 AAAATGAGCAGTAAAATAGATGAGATAAAGAACTAAAGAAAT 5234
            ||| |||||||||||| |||| | || ||||| || | ||
Db    420 CGAATTAGCAGTAAAATACATGATACAAGGAACTCAAAAGAT 461


RESULT   9
AL045429
LOCUS       AL045429      504 bp    mRNA            EST       29-FEB-2000
DEFINITION  DKFZp434F135_r1 434 (synonym: htes3) Homo sapiens cDNA clone
            DKFZp434F135 5', mRNA sequence.
ACCESSION   AL045429
VERSION     AL045429.1  GI:5433574
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 504)
  AUTHORS   Duesterhoeft,A., Lauber,J., Mewes,H.W., Gassenhuber,J. and
            Wiemann,S.
  TITLE     EST (Duesterhoeft, et al.)
  JOURNAL   Unpublished (1999)
COMMENT     On May 5, 1999 this sequence version replaced gi:4753470.
            Contact: Duesterhoeft A
            MIPS
            Am Klopferspitz 18a D-82152 Martinsried, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
            consortium of the German Genome Project.
            No s1 sequence available.
            This clone (DKFZp434F135) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin- Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .504
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="DKFZp434F135"
                     /clone_lib="434 (synonym: htes3)"
                     /tissue_type="testis"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT      116 a     99 c    138 g    146 t      5 others
ORIGIN


  Query Match             5.0%;  Score 285;  DB 47;  Length 504;
  Best Local Similarity   80.5%;  Pred. No. 1.7e-60;
  Matches  408;  Conservative    0;  Mismatches   84;  Indels   15;  Gaps    6;

Qy   4600 CCAAGATTTGGTGCAGGAAAAGTACCTGGAGTA-CGGCAGGTGCCGGACAGTGATCCCGC 4658
```

```
Db      1 CCAAGATTGGGTGCAGGAAAACTACCTGGAGTANCGGCAGGTACCCGGCAGTANTCCTGC 60

Qy   4659 ACGCTATGAGTTCCTGTGGGGTCCAAGGGCCCTCGCTGAAACCAGCTATGTGAAAGTCCT 4718

Db     61 GCGCTATGAGTTCCTGTGGGGTCCAAGGGCTCTGGCTGANACCAGCTATGTGAAAGTCCT 120

Qy   4719 TGAGTATGTGATCAAGGTCAGTGCAAGAGTTCGCTTTTTCTTCCCATCCCTGCGTGAAGC 4778

Db    121 GGAGCATGTGGTCAGGGTCAATGCAAGAGTTCGCATTGCCTACCCATCCCTGCNCGAAGC 180

Qy   4779 AGCTTTGAGAGAGGAGGAAGAGGGAGTCTGAGCATGAGTTGCAGCCAAGGCCAGTGGGAG 4838

Db    181 AGCTTTGTTAGAGGAGGAAGAGGGAGTCTGAGCATGAGTTGCAGCCAGGGCTGTGGGGAA 240

Qy   4839 GG----GGACTGGGCCAGTGCACCTTCCAGGGCCGCGTCCAGCAGCTTCCCCTGCCTCGT 4894

Db    241 GGGGCAGGGCTGGGCCAGTGCATCTAACA--GCCCTGTGCAGCAGCTTCCCTTGCCTCGT 298

Qy   4895 GTGACATGAGGCCCATTCTTCACTCTGAAGAGAGCGGTCAGT------GTTCTCAGTAGT 4948

Db    299 GTAACATGAGGCCCATTCTTCACTCTGTTTGAAGAAAATAGTCAAGTGTTCTTAAGTAGT 358

Qy   4949 AGGTTTCTGTTCTATTGGG-TGACTTGGAGATTTATCTTTGTTCTCTTTTGGAATTGTTC 5007

Db    359 GGGTTTCTATTTTGTTGGGATGACTTGGAGATTTATCTCTGTTTCCTTTTACAATTGTTG 418

Qy   5008 AAATGTTTTTTTTTAAGGGATGGTTGAATGAACTTCAGCATCCAAGTTTATGAATGACAG 5067

Db    419 AAATG-TTCCTTTTAATGGATGGTTGAATTAACTTCAGCATCCAAGTTTATGAATCGTAN 477

Qy   5068 CAGTCACACAGTTCTGTGTATATAGTT 5094

Db    478 TTAACGTATATTGCTGTTAATATAATT 504


RESULT 10
AI863022/c
LOCUS       AI863022      493 bp    mRNA            EST       07-MAR-2000
DEFINITION  wm45d10.x1 NCI_CGAP_Ut4 Homo sapiens cDNA clone IMAGE:2438899 3',
            mRNA sequence.
ACCESSION   AI863022
VERSION     AI863022.1  GI:5527129
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 493)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     On Oct 30, 1998 this sequence version replaced gi:3815061.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1787   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 395.
FEATURES             Location/Qualifiers
     source          1. .493
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2438899"
                     /clone_lib="NCI_CGAP_Ut4"
                     /tissue_type="serous papillary carcinoma, high grade, 2
                     pooled tumors"
                     /lab_host="DH10B"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.48 kb. Life Technologies catalog #:
                     11542-016"
BASE COUNT      184 a     92 c     66 g    151 t
ORIGIN


  Query Match             4.9%;  Score 275.2;  DB 45;  Length 493;
  Best Local Similarity   82.0%;  Pred. No. 4.8e-58;
  Matches  410;  Conservative    0;  Mismatches   63;  Indels   27;  Gaps    7;

Qy   4879 GCTTCCCCTGCCTCGTGTGACATGAGGCCCATTCTTCACTC----TGAAGAGAGCGGTCA 4934

Db    493 GCTTCCCTTGCCTCGTGTAACATGAGGCCCATTCTTCACTCTGTTTGAAGAAAATAGTCA 434

Qy   4935 GTGTTCTCAGTAGTAGGTTTCTGTTCTATTGGGTGACTTGGAGATTTATCTTTGTTCTCT 4994

Db    433 GTGTTCTTAGTAGTGGGTTTCTATTTTGTTGGATGACTTGGAGATTTATCTCTGTTTCCT 374

Qy   4995 TTTGGAATTGTTCAAATGTTTTTTTTTAAGGGATGGTTGAATGAACTTCAGCATCCAAGT 5054

Db    373 TTTACAATTGTTGAAATG-TTCCTTTTAATGGATGGTTGAATTAACTTCAGCATCCAAGT 315

Qy   5055 TTATGAATGACAGCAGTCACACAGTTCTGTGTATATAGTTTAAGGGTAAGAGTCTTGTGT 5114

Db    314 TTATGAATCGTAGTTAACGTATATTGCTGTTAATATAGTTTAGGAGTAAGAGTCTTGTTT 255

Qy   5115 TTTATTCAGATTGGGAAATCCATTCTATTTTGTGAATTGGG---ATAATAACAGCAGTGG 5171

Db    254 TTTATTCAGATTGGGAAATCCGTTCTATTTTGTGAATTTGGGACATAATAACAGCAGTGG 195

Qy   5172 AATAAGTACTTAGAAATGTGAAAAATGAGCAGTAAAATAGATGAGATAAAGAACTAAAGA 5231

Db    194 AGTAAGTATTTAGAAGTGTG---AATTCACCGTGAAATAGGTGAGAT-----------A 150

Qy   5232 AATTAAGAGATAGTCAATTCTTGCCTTATACCTCAGTCTATTCTGTAAAATTT-TTAAAG 5290

Db    149 AATTAAAAGATACTTAATTCCCGCCTTATGCCTCAGTCTATTCTGTAAAATTTAAAAAAT 90

Qy   5291 ATATATGCATACCTGGATTTCCTTGGCTTCTTTGAGAATGTAAGAGAAATTAAATCTGAA 5350

Db     89 ATATATGCATACCTGGATTTCCTTGGCTTC---GTGAATGTAAGAGAAATTAAATCTGAA 33

Qy   5351 TAAAGAATTCTTCCTGTTCA 5370

Db     32 TAAATAATTCTTTCTGTTAA 13


RESULT 11
AW245614/c
LOCUS       AW245614      511 bp    mRNA            EST       07-JAN-2000
DEFINITION  2822978.3prime NIH_MGC_7 Homo sapiens cDNA clone IMAGE:2822978 3',
            mRNA sequence.
ACCESSION   AW245614
VERSION     AW245614.1  GI:6588607
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 511)
  AUTHORS   NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Other_ESTs: 2822978.5prime
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
```

```
            Tissue Procurement: DCTD/DTP cDNA Library Preparation: Ling
            Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E.
            Consortium (LLNL) DNA Sequencing by: Berkeley MGC sequencing
            project Clone distribution: MGC clone distribution information can
            be found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html Base Calling / Quality
            Scores: PHRED from University of Washington Genome Center. Vector
            Trimming: cross_match from University of Washington Genome Center
            PHRAP suite. Poly-T Identification: patMatch.pl from Berkeley
            Drosophila Genome Project. University of Washington Genome Center:
            http://www.genome.washington.edu Polyadenylation: Based upon the
            presence of a XhoI site followed by a run of 14 or more T residues
            at the beginning of the sequence, this cDNA insert was
            polyadenylated.
            Plate: LLCM10  row: N  column: 3
            High quality sequence stop: 427.
FEATURES             Location/Qualifiers
     source          1. .511
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2822978"
                     /clone_lib="NIH_MGC_7"
                     /tissue_type="small cell carcinoma"
                     /cell_line="MGC3"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      174 a    103 c     79 g    155 t
ORIGIN


  Query Match             4.8%;  Score 272.6;  DB 69;  Length 511;
  Best Local Similarity   81.7%;  Pred. No. 2.2e-57;
  Matches  416;  Conservative    0;  Mismatches   74;  Indels   19;  Gaps    8;

Qy    4862 CAGGGCCGCGTCCAGCAGCTTCCCCTGCCTCGTGTGACATGAGGCCCATTCTTCACTC-- 4919
           | ||||||| |||    || | || ||||||| |||||| ||||||||||||||||||
Db     511 CGGGGCCGCATCCCTTAGTTGCCACTGCCTCCTGTGACGTGAGGCCCATTCTTCACTCTT 452

Qy    4920 TGAAGAGAGCGGTCAGTGTTCTCAGTAGTAGGTTTCTGTTCTATTGGGTGACTTGGAGAT 4979
           ||||| |||| |||||  |||| |||||| ||||| |||| | |||| ||||||  |||||
Db     451 TGAAGCGAGCAGTCAGCATTCTTAGTAGTGGGTTTTTGTTTTGTTGGATGACTTTGAGAT 392

Qy    4980 TTATCTTTGTTCTCTTTTGGAATTGTTCAAATGTTTTTTTTTAAGGGATGGTTGAATGAA 5039
           |  ||||||||  || ||||| |||||||||||  ||  |||||| ||||||||||||||
Db     391 TATTCTTTGTTTCCTGTTGGAGTTGTTCAAATG-TTCCTTTTAACGGATGGTTGAATGAG 333

Qy    5040 CTTCAGCATCCAAGTTTATGAATGACAGCAGTCACAC--AGTTCTGTGTATATAGTTTAA 5097
           | |||||||||| |||||||||||||| ||||||||  ||| |||| ||||||||||
Db     332 CGTCAGCATCCAGGTTTATGAATGACAGTAGTCACACATAGTGCTGTTTATATAGTTTAG 273

Qy    5098 GGGTAAGAGTCTTGTGTTTTATTCAGATTGGGAAATCCATTCTATTTTGTGAATTGGG-- 5155
           | |||||||||||||| ||||| ||| |||||||||||||| | ||||||||||||| |
Db     272 GAGTAAGAGTCTTGTTTTTTACTCAAATTGGGAAATCCATCCCATTTTGTGAATTGTGAC 213

Qy    5156 ATAATAACAGCAGTGGAATAAGTA----CTTAGAAATGTGA-AAAATGAGCAGTAAAATA 5210
           ||||||| |||||||| | |||||     |||| || |||||   ||| |||| ||| |||
Db     212 ATAATAATAGCAGTGGTAAAAGTATTTGCTTAAAATTGTGAGCGAATTAGCAATAACATA 153

Qy    5211 GATGAGATAAAGAACTAAAGAAATTAAGAGATAGTCAATTCTTGCCTTATACCTCAGTCT 5270
           |||||||    |||| |||||||| || ||||||||  ||| ||||||| ||||||| | |
Db     152 CATGAGAT----AACTCAAGAAATCAAAAGATAGTTGATTTTTGCCTTGTACCTCAATAT 97

Qy    5271 ATTCTGTAAAATTTTTAAAGATATATGCATACCTGGATTTCCTTGGCTTCTTTGAGAATG 5330
           ||| |||||||   |||| | ||||||| ||| |||||||||| ||||||||||||||
Db      96 ATTGTGTAAAA---TTAAACAAATATGCAAACCAGGATTTCCTTGACTTCTTTGAGAATG 40
```

```
Qy    5331 TAAGAGAAATTAAATCTGAATAAAGAATT 5359
           ||| |||||||||||||| |||| ||||
Db      39 CAAGCGAAATTAAATCTGAGTAAATAATT 11


RESULT  12
AA995045/c
LOCUS       AA995045       481 bp    mRNA            EST       27-AUG-1998
DEFINITION  ou53e02.s1 NCI_CGAP_Br2 Homo sapiens cDNA clone IMAGE:1631546 3',
            mRNA sequence.
ACCESSION   AA995045
VERSION     AA995045.1  GI:3181534
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 481)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     On May 18, 1998 this sequence version replaced gi:3138536.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1718   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 446.
FEATURES             Location/Qualifiers
     source          1. .481
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1631546"
                     /clone_lib="NCI_CGAP_Br2"
                     /sex="female, pooled"
                     /tissue_type="breast"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; 1st strand cDNA was prepared from pooled bulk
                     breast tumor tissue, and was then primed with a Not I -
                     oligo(dT) primer. Double-stranded cDNA was ligated to Eco
                     RI adaptors (Pharmacia), digested with Not I and cloned
                     into the Not I and Eco RI sites of the modified pT7T3
                     vector.  This library is the normalized version of
                     NCI_CGAP_Br1.1. Library was constructed by Bento Soares
                     and M. Fatima Bonaldo. "
BASE COUNT      165 a     91 c     71 g    154 t
ORIGIN


  Query Match             4.8%;  Score 270.6;  DB 33;  Length 481;
  Best Local Similarity   84.6%;  Pred. No. 6.8e-57;
  Matches  402;  Conservative    0;  Mismatches   54;  Indels   19;  Gaps    8;

Qy    4902 GAGGCCCATTCTTCACTC--TGAAGAGAGCGGTCAGTGTTCTCAGTAGTAGGTTTCTGTT 4959
           ||||||||||||||||||  ||||| |||| |||||  |||| |||||||| |||||||||
Db     481 GAGGCCCATTCTTCACTCTTTGAAGCGAGCAGTCAGCATTCTTAGTAGTGGGTTTCTGTT 422

Qy    4960 CTATTGGGTGACTTGGAGATTTATCTTTGTTCTCTTTTGGAATTGTTCAAATGTTTTTTT 5019
           || |||| |||||| ||||||  |||||||  || ||||| |||||||||| ||  ||
Db     421 CTGTTGGATGACTTTGAGATTATTCTTTGTTTCCTGTTGGAGTTGTTCAAATG-TTCCTT 363

Qy    5020 TTAAGGGATGGTTGAATGAACTTCAGCATCCAAGTTTATGAATGACAGCAGTCACAC--A 5077
           |||| |||||||||||||| | |||||||||| ||||||||||||||| ||||||||  |
```

```
Db    362 TTAACGGATGGTTGAATGAGCGTCAGCATCCAGGTTTATGAATGACAGTAGTCACACATA 303

Qy   5078 GTTCTGTGTATATAGTTTAAGGGTAAGAGTCTTGTGTTTTATTCAGATTGGGAAATCCAT 5137
          || |||| |||||||||| | ||||||||||||| ||||||||||||||||||||||||
Db    302 GTGCTGTTTATATAGTTTAGGAGTAAGAGTCTTGTTTTTTATTCAGATTGGGAAATCCAT 243

Qy   5138 TCTATTTTGTGAATTGGG--ATAATAACAGCAGTGGAATAAGTA----CTTAGAAATGTG 5191
          || |||||||||||||  |  ||||||| |||||||||| |||||    |||| || ||||
Db    242 TCCATTTTGTGAATTGTGACATAATAATAGCAGTGGAAAAAGTATTTGCTTAAAATTGTG 183

Qy   5192 A-AAAATGAGCAGTAAAATAGATGAGATAAAGAACTAAAGAAATTAAGAGATAGTCAATT 5250
          |   ||| |||| ||| ||| |||||||    |||| ||||||| || |||||||  |||
Db    182 AGCGAATTAGCAATAACATACATGAGAT----AACTCAAGAAATCAAAAGATAGTTGATT 127

Qy   5251 CTTGCCTTATACCTCAGTCTATTCTGTAAAATTTTTAAAGATATATGCATACCTGGATTT 5310
          |||||||| ||||||| ||||||||||||||   ||||| | ||||||| ||| ||||||
Db    126 CTTGCCTTGTACCTCAATCTATTCTGTAAAA---TTAAACAAATATGCAAACCAGGATTT 70

Qy   5311 CCTTGGCTTCTTTGAGAATGTAAGAGAAATTAAATCTGAATAAAGAATTCTTCCT 5365
          ||||| |||||||||||||| ||| ||||||||||||||||||| ||||||||||
Db     69 CCTTGACTTCTTTGAGAATGCAAGCGAAATTAAATCTGAATAAATAATTCTTCCT 15


RESULT  13
AW194089/c
LOCUS       AW194089      495 bp    mRNA            EST       29-NOV-1999
DEFINITION  xm12a03.x1 NCI_CGAP_Ut4 Homo sapiens cDNA clone IMAGE:2683948 3',
            mRNA sequence.
ACCESSION   AW194089
VERSION     AW194089.1  GI:6472822
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 495)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     On Oct 30, 1998 this sequence version replaced gi:3817926.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Possible reversed clone: polyT not found
            Seq primer: -40UP from Gibco
            High quality sequence stop: 408.
FEATURES             Location/Qualifiers
     source          1. .495
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2683948"
                     /clone_lib="NCI_CGAP_Ut4"
                     /tissue_type="serous papillary carcinoma, high grade, 2
                     pooled tumors"
                     /lab_host="DH10B"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.48 kb. Life Technologies catalog #:
                     11542-016"
BASE COUNT      186 a     95 c     68 g    146 t
ORIGIN
```

```
  Query Match             4.8%;  Score 270.6;  DB 69;  Length 495;
  Best Local Similarity   81.9%;  Pred. No. 6.8e-57;
  Matches  417;  Conservative    0;  Mismatches   64;  Indels   28;  Gaps    8;

Qy   4871 GTCCAGCAGCTTCCCCTGCCTCGTGTGACATGAGGCCCATTCTTCACTC----TGAAGAG 4926
          || |||||||||||| |||||||||| |||||||||||||||||||||    ||||||
Db    494 GTGCAGCAGCTTCCCTTGCCTCGTGTAACATGAGGCCCATTCTTCACTCTGTTTGAAGAA 435

Qy   4927 AGCGGTCAGTGTTCTCAGTAGTAGGTTTCTGTTCTATTGGGTGACTTGGAGATTTATCTT 4986
          |   |||||||||| |||||| ||||||| || | |||| |||||||||||||||||||
Db    434 AATAGTCAGTGTTCTTAGTAGTGGGTTTCTATTTTGTTGGATGACTTGGAGATTTATCTC 375

Qy   4987 TGTTCTCTTTTGGAATTGTTCAAATGTTTTTTTTTAAGGGATGGTTGAATGAACTTCAGC 5046
          ||||  |||||  ||||||| ||||| ||  |||||| |||||||||||| |||||||||
Db    374 TGTTTCCTTTTACAATTGTTGAAATG-TTCCTTTTAATGGATGGTTGAATTAACTTCAGC 316

Qy   5047 ATCCAAGTTT-ATGAATGACAGCAGTCACACAGTTCTGTGTATATAGTTTAAGGGTAAGA 5105
          |||||||||| ||||||   ||     |  | | | |||| |||||||||| | ||||||
Db    315 ATCCAAGTTTAATGAATCGTAGTTAACGTATATTGCTGTTAATATAGTTTAGGAGTAAGA 256

Qy   5106 GTCTTGTGTTTTATTCAGATTGGGAAATCCATTCTATTTTGTGAATTGGG---ATAATAA 5162
          ||||||| |||||||||||||||||||||| |||||||||||||||| ||   |||||||
Db    255 GTCTTGTTTTTTATTCAGATTGGGAAATCCGTTCTATTTTGTGAATTTGGGACATAATAA 196

Qy   5163 CAGCAGTGGAATAAGTACTTAGAAATGTGAAAAATGAGCAGTAAAATAGATGAGATAAAG 5222
          |||||||||| |||||| |||||| |||   |||   | || |||||| ||||||
Db    195 CAGCAGTGGAGTAAGTATTTAGAAGTGTG---AATTCACCGTGAAATAGGTGAGAT---- 141

Qy   5223 AACTAAAGAAATTAAGAGATAGTCAATTCTTGCCTTATACCTCAGTCTATTCTGTAAAAT 5282
                |||||| ||||| | |||||  |||||| ||||||||||||||||||||||||
Db    142 --------AAATTAAAAGATACTTAATTCCCGCCTTATGCCTCAGTCTATTCTGTAAAAT 91

Qy   5283 TT-TTAAAGATATATGCATACCTGGATTTCCTTGGCTTCTTTGAGAATGTAAGAGAAATT 5341
          ||   ||| |||||||||||||||||||||||||||   | ||||||||||||||||
Db     90 TTAAAAAATATATATGCATACCTGGATTTCCTTGGCTTC---GTGAATGTAAGAGAAATT 34

Qy   5342 AAATCTGAATAAAGAATTCTTCCTGTTCA 5370
          ||||||||||||| ||||||| ||||| |
Db     33 AAATCTGAATAAATAATTCTTTCTGTTAA 5


RESULT  14
AW250219/c
LOCUS       AW250219      521 bp    mRNA            EST       07-JAN-2000
DEFINITION  2822505.3prime NIH_MGC_7 Homo sapiens cDNA clone IMAGE:2822505 3',
            mRNA sequence.
ACCESSION   AW250219
VERSION     AW250219.1  GI:6593212
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 521)
  AUTHORS   NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     On Jun 15, 1998 this sequence version replaced gi:3222620.
            Other_ESTs: 2822505.5prime
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: DCTD/DTP cDNA Library Preparation: Ling
            Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E.
            Consortium (LLNL) DNA Sequencing by: Berkeley MGC sequencing
            project Clone distribution: MGC clone distribution information can
            be found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html Base Calling / Quality
            Scores: PHRED from University of Washington Genome Center. Vector
            Trimming: cross_match from University of Washington Genome Center
            PHRAP suite. Poly-T Identification: patMatch.pl from Berkeley
            Drosophila Genome Project. University of Washington Genome Center:
            http://www.genome.washington.edu Polyadenylation: Based upon the
```

presence of a XhoI site followed by a run of 14 or more T residues at the beginning of the sequence, this cDNA insert was polyadenylated.
Plate: LLCM9 row: J column: 10
High quality sequence stop: 453.

FEATURES Location/Qualifiers
source 1. .521
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2822505"
/clone_lib="NIH_MGC_7"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 184 a 100 c 87 g 150 t
ORIGIN

Query Match 4.6%; Score 260.2; DB 70; Length 521;
Best Local Similarity 81.9%; Pred. No. 2.8e-54;
Matches 425; Conservative 0; Mismatches 73; Indels 21; Gaps 10;

```
Qy   4866 GCCGCGTCCAGCAGCTTCCCCTGCCTCGTGTGACATGAGGCCCATTCTTCACTC--TGAA 4923
          ||| | |||   || |||| ||||||| |||||| ||||||||||||||||||||  ||||
Db    521 GCCCCATCCCTTAGTTTCCACTGCCTCCTGTGACGTGAGGCCCATTCTTCACTCTTTGAA 462

Qy   4924 GAGAGCGGTCAGTGTTCTCAGTAGTAGGTTTCTGTTCTATTGGGTGACTTGGAGATTTAT 4983
          | |||| |||||  |||| |||||| |||||||||||| |||| |||||| ||||||  |
Db    461 GCGAGCAGTCAGCATTCTTAGTAGTGGGTTTCTGTTCTGTTGGATGACTTTGAGATTATT 402

Qy   4984 CTTTGTTCTCTTTTGGAATTGTTCAAATGTTTTTTTTTAAGGGATGGTTGAATGAACTTC 5043
          |||||||  || ||||| |||||||||| ||  |||||| |||||||||||||| | ||
Db    401 CTTTGTTTCCTGTTGGAGTTGTTCAAATG-TTCCTTTTAACGGATGGTTGAATGAGCGTC 343

Qy   5044 AGCATCCAAGTTTATGAATGACAGCAGTCACAC--AGTTCTGTGTATATAGTTTAAGGGT 5101
          |||||||| |||||||||||||| ||||||||  ||| |||| |||||||||| | ||
Db    342 AGCATCCAGGTTTATGAATGACAGTAGTCACACATAGTGCTGTTTATATAGTTTAGGAGT 283

Qy   5102 AAGAGTCTTGTGTTTTATTCAGATTGGGAAATCCATTCTATTTTGTGAATTGGG--ATAA 5159
          |||||||||| |||||||||||||||||||||||||| |||||||||||| |  ||||
Db    282 AAGAGTCTTGTTTTTTATTCAGATTGGGAAATCCATTCCATTTTGTGAATTGTGACATAA 223

Qy   5160 TAACAGCAGTGGAATAAGTA----CTTAGAAATGTGAAA-AATGAGCAGTAAAATAGATG 5214
          ||| |||||||||| |||||    |||| || |||||   ||| |||| ||| ||| |||
Db    222 TAATAGCAGTGGAAAAAGTATTTGCTTAAAATTGTGAGCGAATTAGCAATAACATACATG 163

Qy   5215 AGATAAAGAACTAAAGAAATTAAGAGATAGTCAATTCTTG-CCTTATACCTCAGTCTATT 5273
          ||||    |||| ||||||| || |||||||  |||||| |||| ||||||| ||||||
Db    162 AGAT----AACTCAAGAAATCAAAAGATAGTTGATTCTTGCCCTTGTACCTCAATCTATT 107

Qy   5274 CTGTAAAATTTTTAAAGATATATGCATACCTGGATTTCCTTGGCTTCTTTGAGAATGTAA 5333
          ||||||||   ||||| | ||||||| ||| |||||||||| |||||||| || || ||
Db    106 CTGTAAAA---TTAAACAAATATGCAAACCAGGATTTCCTTGACTTCTTTGGGATTGCAA 50

Qy   5334 -GAGAAATTAAATCTGAATAAAGAATTCTTCCTGTTCAC 5371
           | |||||| ||| || || ||  ||| ||||| ||| |
Db     49 GGCGAAATTTAATTTGGATTAAATATTTTTCCTTTTCCC 11
```

RESULT 15
AI222439/c
LOCUS AI222439 459 bp mRNA EST 30-NOV-1998
DEFINITION qh04h04.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:1843735 3', mRNA sequence.
ACCESSION AI222439
VERSION AI222439.1 GI:3804642
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 459)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 770 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 439.

FEATURES Location/Qualifiers
source 1. .459
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1843735"
/clone_lib="Soares_NFL_T_GBC_S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NbHL19W, testis NHT, and B-cell NCI_CGAP_GCB1) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 159 a 84 c 65 g 151 t
ORIGIN

Query Match 4.5%; Score 257.8; DB 36; Length 459;
Best Local Similarity 83.7%; Pred. No. 1.1e-53;
Matches 379; Conservative 0; Mismatches 57; Indels 17; Gaps 7;

```
Qy   4927 AGCGGTCAGTGTTCTCAGTAGTAGGTTTCTGTTCTATTGGGTGACTTGGAGATTTATCTT 4986
          ||| |||||  |||| |||||| |||||||||||| |||| |||||| ||||||  ||||
Db    459 AGCAGTCAGCATTCTTAGTAGTGGGTTTCTGTTCTGTTGGATGACTTTGAGATTATTCTT 400

Qy   4987 TGTTCTCTTTTGGAATTGTTCAAATGTTTTTTTTTAAGGGATGGTTGAATGAACTTCAGC 5046
          ||||  || ||||| |||||||||| ||  |||||| |||||||||||||| | |||||
Db    399 TGTTTCCTGTTGGAGTTGTTCAAATG-TTCCTTTTAACGGATGGTTGAATGAGCGTCAGC 341

Qy   5047 ATCCAAGTTTATGAATGACAGCAGTCACAC--AGTTCTGTGTATATAGTTTAAGGGTAAG 5104
          ||||| |||||||||||||||| ||||||||  ||| |||| |||||||||| | |||||
Db    340 ATCCAGGTTTATGAATGACAGTAGTCACACATAGTGCTGTTTATATAGTTTAGGAGTAAG 281

Qy   5105 AGTCTTGTGTTTTATTCAGATTGGGAAATCCATTCTATTTTGTGAATTGGG--ATAATAA 5162
          |||||||| ||||| ||| |||||||||||||||| |||||||||||| |  |||||||
Db    280 AGTCTTGTTTTTTACTCAAATTGGGAAATCCATTCCATTTTGTGAATTGTGACATAATAA 221

Qy   5163 CAGCAGTGGAATAAGTA----CTTAGAAATGTGA-AAAATGAGCAGTAAAATAGATGAGA 5217
           ||||||||  | |||||    |||| || |||||   ||| |||| ||| ||| ||||||
Db    220 TAGCAGTGGTAAAAGTATTTGCTTAAAATTGTGAGCGAATTAGCAATAACATACATGAGA 161

Qy   5218 TAAAGAACTAAAGAAATTAAGAGATAGTCAATTCTTGCCTTATACCTCAGTCTATTCTGT 5277
          |    |||| ||||||| || ||||||| |||||||||| |||||||| ||||||||||
Db    160 T----AACTCAAGAAATCAAAAGATAGTTGATTCTTGCCTTGTACCTCAATCTATTCTGT 105

Qy   5278 AAAATTTTTAAAGATATATGCATACCTGGATTTCCTTGGCTTCTTTGAGAATGTAAGAGA 5337
          ||||   ||||| | ||||||| ||| |||||||||| |||||||||||||| ||| ||
Db    104 AAAA---TTAAACAAATATGCAAACCAGGATTTCCTTGACTTCTTTGAGAATGCAAGCGA 48
```

```
Qy    5338 AATTAAATCTGAATAAAGAATTCTTCCTGTTCA 5370
           ||||||||||||||||| |||||||||| ||||
Db      47 AATTAAATCTGAATAAATAATTCTTCCTCTTCA 15
```

Search completed: September 13, 2000, 02:38:16
Job time: 4783 sec

****

MPSRCH (TM)

****

Release 3.1A John F. Collins, Biocomputing Research Unit.

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Sep 13 07:14:42 2000; MasPar time 3.43 Seconds
40.218 Million cell updates/sec

Tabular output not generated.

Title: >US-08-819-669E-26
Description: (1-9) from US08819669E.pep
Perfect Score: 61
Sequence: 1 EADPTGHSY 9

Scoring table: PAM 150
Gap 15

Searched: 152433 seqs, 15329240 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-issued
1:5A_COMB 2:5B_COMB 3:6_COMB 4:PCT_COMB 5:backfiles1

Statistics: Mean 14.653; Variance 34.108; scale 0.430

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	61	100.0	9	2	US-09-036-	Sequence 1, Applicatio	1.24e-01
2	61	100.0	9	2	US-08-986-	Sequence 1, Applicatio	1.24e-01
3	61	100.0	9	1	US-07-938-	Sequence 1, Applicatio	1.24e-01
4	61	100.0	9	1	US-08-443-	Sequence 12, Applicati	1.24e-01
5	61	100.0	9	1	US-08-787-	Sequence 49, Applicati	1.24e-01
6	61	100.0	9	1	US-08-073-	Sequence 12, Applicati	1.24e-01
7	61	100.0	9	3	US-08-159-	Sequence 99, Applicati	1.24e-01
8	61	100.0	9	3	US-08-354-	Sequence 12, Applicati	1.24e-01
9	61	100.0	9	2	US-08-902-	Sequence 21, Applicati	1.24e-01
10	61	100.0	9	2	US-08-142-	Sequence 26, Applicati	1.24e-01
11	61	100.0	9	4	PCT-US95-0	Sequence 2, Applicatio	1.24e-01
12	61	100.0	9	4	PCT-US95-0	Sequence 1, Applicatio	1.24e-01
13	61	100.0	9	1	US-08-299-	Sequence 26, Applicati	1.24e-01
14	61	100.0	9	2	US-08-498-	Sequence 4, Applicatio	1.24e-01
15	61	100.0	9	3	US-08-967-	Sequence 26, Applicati	1.24e-01
16	61	100.0	9	1	US-08-186-	Sequence 1, Applicatio	1.24e-01
17	61	100.0	10	3	US-08-602-	Sequence 25, Applicati	1.24e-01
18	61	100.0	10	1	US-08-796-	Sequence 25, Applicati	1.24e-01
19	61	100.0	10	2	US-08-498-	Sequence 5, Applicatio	1.24e-01
20	61	100.0	12	2	US-08-560-	Sequence 4, Applicatio	1.24e-01
21	61	100.0	12	1	US-08-190-	Sequence 4, Applicatio	1.24e-01
22	61	100.0	309	1	US-08-465-	Sequence 24, Applicati	1.24e-01
23	61	100.0	309	2	US-08-993-	Sequence 10, Applicati	1.24e-01
24	61	100.0	309	3	US-08-845-	Sequence 10, Applicati	1.24e-01
25	59	96.7	9	3	US-08-159-	Sequence 1201, Applica	2.48e-01
26	57	93.4	9	3	US-08-159-	Sequence 1196, Applica	4.92e-01
27	55	90.2	8	1	US-08-073-	Sequence 14, Applicati	9.70e-01
28	55	90.2	8	1	US-08-443-	Sequence 14, Applicati	9.70e-01
29	55	90.2	8	3	US-08-354-	Sequence 14, Applicati	9.70e-01
30	55	90.2	9	2	US-08-498-	Sequence 8, Applicatio	9.70e-01
31	53	86.9	9	3	US-08-159-	Sequence 1193, Applica	1.90e+00
32	53	86.9	9	3	US-08-159-	Sequence 1198, Applica	1.90e+00
33	52	85.2	9	2	US-08-498-	Sequence 1, Applicatio	2.65e+00
34	52	85.2	9	2	US-08-498-	Sequence 6, Applicatio	2.65e+00
35	52	85.2	369	2	US-08-773-	Sequence 4, Applicatio	2.65e+00
36	51	83.6	9	3	US-08-159-	Sequence 1200, Applica	3.70e+00
37	49	80.3	8	1	US-08-073-	Sequence 13, Applicati	7.15e+00
38	49	80.3	8	3	US-08-354-	Sequence 13, Applicati	7.15e+00
39	49	80.3	8	1	US-08-443-	Sequence 13, Applicati	7.15e+00
40	49	80.3	8	1	US-07-938-	Sequence 21, Applicati	7.15e+00
41	49	80.3	9	3	US-08-159-	Sequence 1195, Applica	7.15e+00
42	48	78.7	873	3	US-09-187-	Sequence 6, Applicatio	9.91e+00
43	48	78.7	925	2	US-08-504-	Sequence 1, Applicatio	9.91e+00
44	48	78.7	925	2	US-08-392-	Sequence 1, Applicatio	9.91e+00
45	48	78.7	925	4	PCT-US94-1	Sequence 1, Applicatio	9.91e+00

ALIGNMENTS

RESULT 1
ID US-09-036-582-1 STANDARD; PRT; 9 AA.
XX
AC xxxxxx
XX
DT
XX
DE Sequence 1, Application US/09036582A
XX
CC Sequence 1, Application US/09036582A
CC Patent No. 5965381
CC GENERAL INFORMATION:
CC APPLICANT: van der Bruggen, Pierre
CC APPLICANT: Cornelis, Guy R.
CC TITLE OF INVENTION: DELIVERY OF PROTEINS INTO EUKARYOTIC CELLS
CC TITLE OF INVENTION: WITH RECOMBINANT YERSINIA
CC FILE REFERENCE: 11154
CC CURRENT APPLICATION NUMBER: US/09/036,582A
CC CURRENT FILING DATE: 1998-03-06
CC NUMBER OF SEQ ID NOS: 39
CC SOFTWARE: PatentIn Ver. 2.0
CC SEQ ID NO 1
CC LENGTH: 9
CC TYPE: PRT
CC ORGANISM: Human MAGE-1 peptide
SQ SEQUENCE 9 AA; 976 MW; 576 CN;

Query Match 100.0%; Score 61; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.24e-01;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 EADPTGHSY 9
|||||||||
Qy 1 EADPTGHSY 9

RESULT 2
ID US-08-986-234-1 STANDARD; PRT; 9 AA.
XX
AC xxxxxx
XX
DT
XX
DE Sequence 1, Application US/08986234
XX
CC Sequence 1, Application US/08986234
CC Patent No. 5981706

```
CC    GENERAL INFORMATION:
CC     APPLICANT: Wallen, et al.
CC     TITLE OF INVENTION: Methods for Synthesizing Heat Shock Protein Complexes
CC     FILE REFERENCE: UNME-0008-1
CC     CURRENT APPLICATION NUMBER: US/08/986,234
CC     CURRENT FILING DATE: 1997-12-05
CC     NUMBER OF SEQ ID NOS: 114
CC     SOFTWARE: PatentIn Ver. 2.0
CC    SEQ ID NO 1
CC      LENGTH: 9
CC      TYPE: PRT
CC      ORGANISM: human
SQ   SEQUENCE   9 AA; 976 MW; 576 CN;

  Query Match          100.0%;  Score 61;  DB 2;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 1.24e-01;
  Matches     9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT    3
ID   US-07-938-334C-1     STANDARD;     PRT;   9 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 1, Application US/07938334C
XX
CC    Sequence 1, Application US/07938334C
CC    Patent No. 5405940
CC     GENERAL INFORMATION:
CC       APPLICANT:  Boon, Thierry; van der Bruggen, Pierre;
CC       APPLICANT:  De Plaen, Etienne; Lurquin Christophe; Traversari, Catia
CC       TITLE OF INVENTION:  ISOLATED NONAPEPTIDES DERIVED FROM
CC       TITLE OF INVENTION:  MAGE GENES AND USES THEREOF
CC       NUMBER OF SEQUENCES:  22
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         COUNTRY:  USA
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
CC         COMPUTER:  IBM PS/2
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/07/938,334C
CC         FILING DATE:  31-AUG-1992
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hanson, No. 5405940man D.
CC         REGISTRATION NUMBER:  30,946
CC         REFERENCE/DOCKET NUMBER:  LUD 293
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  1:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  9 amino acid residues
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
CC       FEATURE:
CC         NAME/KEY:  MAGE-1 derived nonapeptide
SQ   SEQUENCE   9 AA; 976 MW; 576 CN;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 1.24e-01;
  Matches     9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT    4
ID   US-08-443-341-12     STANDARD;     PRT;   9 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 12, Application US/08443341
XX
CC    Sequence 12, Application US/08443341
CC    Patent No. 5695994
CC     GENERAL INFORMATION:
CC       APPLICANT:  Boon-Falleur, Thierry
CC       APPLICANT:  van der Bruggen, Pierre
CC       APPLICANT:  De Plaen, Etienne
CC       APPLICANT:  Lurquin, Christophe
CC       APPLICANT:  Traversari, Catia
CC       APPLICANT:  Gaugler, Beatrice
CC       APPLICANT:  Van den Eynde, Benoit
CC       TITLE OF INVENTION:  ISOLATED NONAPEPTIDES DERIVED FROM
CC       TITLE OF INVENTION:  MAGE-3 GENE AND PRESENTED BY HLA-A1 AND USES THEREOF
CC       NUMBER OF SEQUENCES:  22
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         COUNTRY:  USA
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
CC         COMPUTER:  IBM PS/2
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/443,341
CC         FILING DATE:  17-MAY-1995
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/073,103
CC         FILING DATE:  7-JUNE-1993
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/938,334
CC         FILING DATE:  31-AUG-1992
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/037,230
CC         FILING DATE:  26-MARCH-1993
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hanson, No. 5695994man D.
CC         REGISTRATION NUMBER:  30,946
CC         REFERENCE/DOCKET NUMBER:  LUD 5293.5
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  12:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  9 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   9 AA; 976 MW; 576 CN;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
```

```
  Best Local Similarity 100.0%;  Pred. No. 1.24e-01;
  Matches     9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db       1 EADPTGHSY 9
           |||||||||
Qy       1 EADPTGHSY 9


RESULT    5
ID   US-08-787-547-49     STANDARD;     PRT;   9 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 49, Application US/08787547
XX
CC    Sequence 49, Application US/08787547
CC    Patent No. 5783567
CC     GENERAL INFORMATION:
CC       APPLICANT:  Hedley, Mary Lynne
CC       APPLICANT:  Curley, Joanne M.
CC       APPLICANT:  Langer, Robert S.
CC       TITLE OF INVENTION:  MICROPARTICLES FOR DELIVERY
CC       TITLE OF INVENTION:  OF NUCLEIC ACID
CC       NUMBER OF SEQUENCES:  107
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Fish & Richardson, P.C.
CC         STREET:  225 Franklin Street
CC         CITY:  Boston
CC         STATE:  MA
CC         COUNTRY:  US
CC         ZIP:  02110-2804
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette
CC         COMPUTER:  IBM Compatible
CC         OPERATING SYSTEM:  Windows95
CC         SOFTWARE:  FastSEQ for Windows Version 2.0
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/787,547
CC         FILING DATE:  22-JAN-1997
CC         CLASSIFICATION:  514
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:
CC         FILING DATE:
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Fraser, Janis K.
CC         REGISTRATION NUMBER:  34,819
CC         REFERENCE/DOCKET NUMBER:  08191/003001
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  617-542-5070
CC         TELEFAX:  617-542-8906
CC         TELEX:  200154
CC     INFORMATION FOR SEQ ID NO:  49:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  9 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  peptide
SQ   SEQUENCE   9 AA; 976 MW; 576 CN;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 1.24e-01;
  Matches     9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db       1 EADPTGHSY 9
           |||||||||
Qy       1 EADPTGHSY 9


RESULT    6
ID   US-08-073-103A-12     STANDARD;     PRT;   9 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 12, Application US/08073103A
XX
CC    Sequence 12, Application US/08073103A
CC    Patent No. 5462871
CC     GENERAL INFORMATION:
CC       APPLICANT:  Boon-Falleur, Thierry
CC       APPLICANT:  van der Bruggen, Pierre
CC       APPLICANT:  De Plaen, Etienne
CC       APPLICANT:  Lurquin, Christophe
CC       APPLICANT:  Traversari, Catia
CC       APPLICANT:  Gaugler, Beatrice
CC       APPLICANT:  Van den Eynde, Benoit
CC       TITLE OF INVENTION:  ISOLATED NONAPEPTIDES DERIVED FROM
CC       TITLE OF INVENTION:  MAGE-3 GENE AND PRESENTED BY HLA-A1 AND USES THEREOF
CC       NUMBER OF SEQUENCES:  22
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         COUNTRY:  USA
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
CC         COMPUTER:  IBM PS/2
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/073,103A
CC         FILING DATE:  7-JUNE-1993
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/938,334
CC         FILING DATE:  31-AUG-1992
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hanson, No. 5462871man D.
CC         REGISTRATION NUMBER:  30,946
CC         REFERENCE/DOCKET NUMBER:  LUD 5293.1
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  12:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  9 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   9 AA; 976 MW; 576 CN;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 1.24e-01;
  Matches     9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db       1 EADPTGHSY 9
           |||||||||
Qy       1 EADPTGHSY 9


RESULT    7
ID   US-08-159-339A-99     STANDARD;     PRT;   9 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 99, Application US/08159339A
XX
CC    Sequence 99, Application US/08159339A
```

```
CC    Patent No. 6037135
CC     GENERAL INFORMATION:
CC       APPLICANT:  Kubo, Ralph T.
CC       APPLICANT:  Grey, Howard M.
CC       APPLICANT:  Sette, Alessandro
CC       APPLICANT:  Celis, Esteban
CC       TITLE OF INVENTION:  HLA Binding peptides and Their
CC       TITLE OF INVENTION:  Uses
CC       NUMBER OF SEQUENCES:  1254
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Townsend and Townsend and Crew LLP
CC         STREET:  Two Embarcadero Center, Eighth Floor
CC         CITY:  San Francisco
CC         STATE:  CA
CC         COUNTRY:  USA
CC         ZIP:  94111-3834
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette
CC         COMPUTER:  IBM Compatible
CC         OPERATING SYSTEM:  DOS
CC         SOFTWARE:  FastSEQ for Windows Version 2.0
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/159,339A
CC         FILING DATE:  29-NOV-1993
CC         CLASSIFICATION:  424
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 07/926,666
CC         FILING DATE:  07-AUG-1992
CC         APPLICATION NUMBER:  US 08/027,746
CC         FILING DATE:  05-MAR-1993
CC         APPLICATION NUMBER:  US 08/103,396
CC         FILING DATE:  06-AUG-1993
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Weber, Ellen Lauver
CC         REGISTRATION NUMBER:  32,762
CC         REFERENCE/DOCKET NUMBER:  018623-005030US
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (415) 576-0200
CC         TELEFAX:  (415) 576-0300
CC         TELEX:
CC     INFORMATION FOR SEQ ID NO:  99:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  9 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  peptide
SQ   SEQUENCE   9 AA; 976 MW; 576 CN;

  Query Match          100.0%;  Score 61;  DB 3;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 1.24e-01;
  Matches     9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT    8
ID   US-08-354-679C-12     STANDARD;     PRT;   9 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 12, Application US/08354679C
XX
CC    Sequence 12, Application US/08354679C
CC    Patent No. 6034214
CC     GENERAL INFORMATION:
CC       APPLICANT:  Boon, Thierry; van der Bruggen, Pierre;
CC       APPLICANT:  De Plaen, Etienne; Lurquin Christophe; Traversari, Catia
CC       TITLE OF INVENTION:  ISOLATED NONAPEPTIDES DERIVED FROM
CC       TITLE OF INVENTION:  MAGE GENES AND USES THEREOF
CC       NUMBER OF SEQUENCES:  25
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         COUNTRY:  USA
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 3.5 inch, 1.44 mb storage
CC         COMPUTER:  IBM PS/2
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/354,679C
CC         FILING DATE:  13-DECEMBER-1994
CC         CLASSIFICATION:  530
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/938,334
CC         FILING DATE:  31-AUGUST-1992
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  BAER, MADELINE F.
CC         REGISTRATION NUMBER:  36,437
CC         REFERENCE/DOCKET NUMBER:  LUD 5293.2
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  12:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  9 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   9 AA; 976 MW; 576 CN;

  Query Match          100.0%;  Score 61;  DB 3;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 1.24e-01;
  Matches     9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT    9
ID   US-08-902-516-21     STANDARD;     PRT;   9 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 21, Application US/08902516
XX
CC    Sequence 21, Application US/08902516
CC    Patent No. 5891432
CC     GENERAL INFORMATION:
CC       APPLICANT:  Soo Hoo, William
CC       TITLE OF INVENTION:  MEMBRANE-BOUND CYTOKINE COMPOSITIONS
CC       TITLE OF INVENTION:  COMPRISING GM-CSF AND METHODS OF MODULATING AN IMMUNE
CC       TITLE OF INVENTION:  RESPONSE USING SAME
CC       NUMBER OF SEQUENCES:  50
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  CAMPBELL & FLORES, LLP
CC         STREET:  4370 La Jolla Village Drive, Suite 700
CC         CITY:  San Diego
CC         STATE:  California
CC         COUNTRY:  United States
CC         ZIP:  92121
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
```

```
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/902,516
CC         FILING DATE:  29-JUL-1997
CC         CLASSIFICATION:  424
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Campbell, Cathryn A.
CC         REGISTRATION NUMBER:  31,815
CC         REFERENCE/DOCKET NUMBER:  P-IM 2442
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (619)535-9001
CC         TELEFAX:  (619)535-8949
CC     INFORMATION FOR SEQ ID NO:  21:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  9 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  peptide
SQ   SEQUENCE   9 AA; 976 MW; 576 CN;

  Query Match          100.0%;  Score 61;  DB 2;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 1.24e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   10
ID   US-08-142-368A-26     STANDARD;     PRT;   9 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 26, Application US/08142368A
XX
CC    Sequence 26, Application US/08142368A
CC    Patent No. 5925729
CC     GENERAL INFORMATION:
CC       APPLICANT:  Boon-Falleur, Thierry; Van der Bruggen, Thierry;
CC       APPLICANT:  Van den Eynde, Beno t; Van Pel, Aline; De Plaen, Etienne;
CC       APPLICANT:  Lurquin, Christophe; Chomez, Patrick; Traversari, Catia
CC       TITLE OF INVENTION:  Tumor Rejection Antigen Precursors, Tumor
CC       TITLE OF INVENTION:  Rejection Antigens and Uses Thereof
CC       NUMBER OF SEQUENCES:  26
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
CC         COMPUTER:  IBM
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/142,368A
CC         FILING DATE:  02-MAY-1994
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US92/04354
CC         FILING DATE:  22-MAY-1992
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/807,043
CC         FILING DATE:  12-DECEMBER-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/764,364
CC         FILING DATE:  23-SEPTEMBER-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/728,838
CC         APPLICATION NUMBER:  9-JULY-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/705,702
CC         FILING DATE:  23-May-1991
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hanson, No. 5925729man D.
CC         REGISTRATION NUMBER:  30,946
CC         REFERENCE/DOCKET NUMBER:  LUD 5253.4-US
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  26:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  9 amino acids
CC         TYPE:  amino acids
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   9 AA; 976 MW; 576 CN;

  Query Match          100.0%;  Score 61;  DB 2;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 1.24e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   11
ID   PCT-US95-04975-2     STANDARD;     PRT;   9 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 2, Application PC/TUS9504975
XX
CC    Sequence 2, Application PC/TUS9504975
CC     GENERAL INFORMATION:
CC       APPLICANT:  Nikolic-Zugic, Janko
CC       APPLICANT:  Dyall, Rubendra
CC       TITLE OF INVENTION:  INDUCTION OF CYTOTOXIC T LYMPHOCYTES (CTL)USING
CC       TITLE OF INVENTION:  ANTIGENIC PEPTIDES AND A SUITABLE ADJUVANT
CC       NUMBER OF SEQUENCES:  19
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Cooper & Dunham LLP
CC         STREET:  1185 Avenue of the Americas
CC         CITY:  New York
CC         STATE:  New York
CC         COUNTRY:  USA
CC         ZIP:  10036
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.24
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US95/04975
CC         FILING DATE:
CC         CLASSIFICATION:
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/233,496
CC         FILING DATE:  April 22, 1994
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  White Esq., John P.
CC         REGISTRATION NUMBER:  28,678
CC         REFERENCE/DOCKET NUMBER:  45059/JPW/MSC/AMB
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  212-278-0400
CC         TELEFAX:  212-391-0525
CC     INFORMATION FOR SEQ ID NO:  2:
CC       SEQUENCE CHARACTERISTICS:
```

```
CC         LENGTH:  9 amino acids
CC         TYPE:  amino acids
CC         TOPOLOGY:  Linear
CC       MOLECULE TYPE:  peptide
CC       HYPOTHETICAL:  N
CC       ANTI-SENSE:  N
SQ   SEQUENCE   9 AA; 976 MW; 576 CN;

  Query Match          100.0%;  Score 61;  DB 4;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 1.24e-01;
  Matches     9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   12
ID   PCT-US95-02121-1     STANDARD;     PRT;   9 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 1, Application PC/TUS9502121
XX
CC    Sequence 1, Application PC/TUS9502121
CC     GENERAL INFORMATION:
CC       APPLICANT:
CC       TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR ELICITING
CC       TITLE OF INVENTION:  CTL IMMUNITY
CC       NUMBER OF SEQUENCES:  153
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US95/02121
CC         FILING DATE:  16-FEB-1995
CC         CLASSIFICATION:
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 08/197,484
CC         FILING DATE:  16-FEB-1994
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 07/935,811
CC         FILING DATE:  26-AUG-1992
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 07/874,491
CC         FILING DATE:  27-APR-1992
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 07/827,682
CC         FILING DATE:  29-JAN-1992
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 07/749,568
CC         FILING DATE:  26-AUG-1991
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Parmelee, Steven W.
CC         REGISTRATION NUMBER:  31,990
CC         REFERENCE/DOCKET NUMBER:  14137-26-4PC
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (206) 467-9600
CC         TELEFAX:  (415) 543-5043
CC     INFORMATION FOR SEQ ID NO:  1:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  9 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  unknown
CC         TOPOLOGY:  unknown
CC       MOLECULE TYPE:  peptide
SQ   SEQUENCE   9 AA; 976 MW; 576 CN;

  Query Match          100.0%;  Score 61;  DB 4;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 1.24e-01;
  Matches     9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   13
ID   US-08-299-849B-26     STANDARD;     PRT;   9 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 26, Application US/08299849B
XX
CC    Sequence 26, Application US/08299849B
CC    Patent No. 5612201
CC     GENERAL INFORMATION:
CC       APPLICANT:  De Plaen, Etienne; Boon-Falleur, Thierry;
CC       APPLICANT:  Leth , Bernard; Szikora, Jean-Pierre; De Smet, Charles;
CC       APPLICANT:  Chomez, Patrick
CC       TITLE OF INVENTION:  Isolated Nucleic Acid Molecules Useful In
CC       TITLE OF INVENTION:  Determining Expression Of A Tumor Antigen Precursor
CC       NUMBER OF SEQUENCES:  48
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
CC         COMPUTER:  IBM
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/299,849B
CC         FILING DATE:  1-SEPTEMBER-1994
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/037,230
CC         FILING DATE:  26-MARCH-1993
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US92/04354
CC         FILING DATE:  22-MAY-1992
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/807,043
CC         FILING DATE:  12-DECEMBER-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/764,364
CC         FILING DATE:  23-SEPTEMBER-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/728,838
CC         APPLICATION NUMBER:  9-JULY-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/705,702
CC         FILING DATE:  23-May-1991
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hanson, No. 5612201man D.
CC         REGISTRATION NUMBER:  30,946
CC         REFERENCE/DOCKET NUMBER:  LUD 5355
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  26:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  9 amino acids
CC         TYPE:  amino acids
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   9 AA; 976 MW; 576 CN;
```

```
  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 1.24e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   14
ID   US-08-498-461-4     STANDARD;     PRT;   9 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 4, Application US/08498461
XX
CC    Sequence 4, Application US/08498461
CC    Patent No. 5827073
CC     GENERAL INFORMATION:
CC       APPLICANT:  Luescher, Immanuel; Anjuere, Fabienee;
CC       APPLICANT:  Layer, Andreas; Romero, Pedro; Cerottini, Jean-Charles
CC       TITLE OF INVENTION:  Photoreactive Peptide Derivatives
CC       NUMBER OF SEQUENCES:  16
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 3.5 inch, 1.44 kb storage
CC         COMPUTER:  IBM
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/498,461
CC         FILING DATE:  5-JULY-1995
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hanson, No. 5827073man D.
CC         REGISTRATION NUMBER:  30,946
CC         REFERENCE/DOCKET NUMBER:  LUD 5403
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  4:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  9 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
SQ   SEQUENCE   9 AA; 976 MW; 576 CN;

  Query Match          100.0%;  Score 61;  DB 2;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 1.24e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   15
ID   US-08-967-727-26     STANDARD;     PRT;   9 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 26, Application US/08967727
XX
CC    Sequence 26, Application US/08967727
CC    Patent No. 6025474
CC     GENERAL INFORMATION:
CC       APPLICANT:  Gaugler, B atrice; Van den Eynde, Beno t;
CC       APPLICANT:  van der Bruggen, Pierre;  Boon-Falleur, Thierry
CC       TITLE OF INVENTION:  Isolated Nucleic Acid Molecules Coding For
CC       TITLE OF INVENTION:  Tumor Rejection Antigen Precursor Mage-3 And Uses There
CC       NUMBER OF SEQUENCES:  30
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
CC         COMPUTER:  IBM
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/967,727
CC         FILING DATE:
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/037,230
CC         FILING DATE:  26-MARCH-1993
CC         APPLICATION NUMBER:  PCT/US92/04354
CC         FILING DATE:  22-MAY-1992
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/807,043
CC         FILING DATE:  12-DECEMBER-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/764,365
CC         FILING DATE:  23-SEPTEMBER-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/728,838
CC         FILING DATE:  9-JULY-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/705,702
CC         FILING DATE:  23-MAY-1991
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hanson, No. 6025474man D.
CC         REGISTRATION NUMBER:  30,946
CC         REFERENCE/DOCKET NUMBER:  LUD 5353
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  26:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  9 amino acids
CC         TYPE:  amino acids
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   9 AA; 976 MW; 576 CN;

  Query Match          100.0%;  Score 61;  DB 3;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 1.24e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   16
ID   US-08-186-266-1     STANDARD;     PRT;   9 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 1, Application US/08186266
XX
```

```
CC    Sequence 1, Application US/08186266
CC    Patent No. 5662907
CC     GENERAL INFORMATION:
CC       APPLICANT:  KUBO, Ralph T.
CC       APPLICANT:  GREY, Howard M.
CC       APPLICANT:  SETTE, Alessandro
CC       APPLICANT:  CELIS, Esteban
CC       TITLE OF INVENTION:  INDUCTION OF ANTI-TUMOR CYTOTOXIC
CC       TITLE OF INVENTION:  T LYMPHOCYTES IN HUMANS USING
CC       TITLE OF INVENTION:  SYNTHETIC PEPTIDE EPITOPES
CC       NUMBER OF SEQUENCES:  20
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Townsend and Townsend Khourie and Crew
CC         STREET:  Steuart Street Tower, One Market Plaza
CC         CITY:  San Francisco
CC         STATE:  California
CC         COUNTRY:  US
CC         ZIP:  94105-1493
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/186,266
CC         FILING DATE:  25-JAN-1994
CC         CLASSIFICATION:  424
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 08/159,339
CC         FILING DATE:  29-NOV-1993
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 08/103,396
CC         FILING DATE:  06-AUG-1993
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 08/027,746
CC         FILING DATE:  05-MAR-1993
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 07/926,666
CC         FILING DATE:  07-AUG-1992
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Bastian, Kevin L.
CC         REGISTRATION NUMBER:  34,774
CC         REFERENCE/DOCKET NUMBER:  14137-50-4
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (415) 543-9600
CC         TELEFAX:  (415) 543-5043
CC     INFORMATION FOR SEQ ID NO:  1:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  9 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  peptide
SQ   SEQUENCE   9 AA; 976 MW; 576 CN;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 1.24e-01;
  Matches     9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   17
ID   US-08-602-506A-25     STANDARD;     PRT;   10 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 25, Application US/08602506A
XX
CC    Sequence 25, Application US/08602506A
CC    Patent No. 6060257
CC     GENERAL INFORMATION:
CC       APPLICANT:  Herman, Jean;  Coulie, Pierre;
CC       APPLICANT:  Boon-Falleur, Thierry; van der Bruggen, Pierre;
CC       APPLICANT:  Luescher, Immanuel.
CC       TITLE OF INVENTION:  Tumor Rejection Antigens Presented By HLA-
CC       TITLE OF INVENTION:  B44 Molecules, And Uses Thereof
CC       NUMBER OF SEQUENCES:  30
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 3.5 inch, 360 kb storage
CC         COMPUTER:  IBM
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/602,506A
CC         FILING DATE:  20-FEBRUARY-1996
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/531,864
CC         FILING DATE:  21-SEPTEMBER-1995
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/373,636
CC         FILING DATE:  17-JANUARY-1995
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/253,503
CC         FILING DATE:  3-JUNE-1994
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hanson, No. 6060257man D.
CC         REGISTRATION NUMBER:  30,946
CC         REFERENCE/DOCKET NUMBER:  LUD 5436
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  25:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  10 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
CC       FEATURE:
CC         NAME/KEY:  MAGE-1/HLA-B44
SQ   SEQUENCE   10 AA; 1104 MW; 684 CN;

  Query Match          100.0%;  Score 61;  DB 3;  Length 10;
  Best Local Similarity 100.0%;  Pred. No. 1.24e-01;
  Matches     9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      2 EADPTGHSY 10
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   18
ID   US-08-796-883-25     STANDARD;     PRT;   10 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 25, Application US/08796883
XX
CC    Sequence 25, Application US/08796883
CC    Patent No. 5744353
CC     GENERAL INFORMATION:
CC       APPLICANT:  Herman, Jean;  Coulie, Pierre;
CC       APPLICANT:  Boon-Falleur, Thierry; van der Bruggen, Pierre;
CC       APPLICANT:  Luescher, Immanuel.
```

```
CC       TITLE OF INVENTION:  Tumor Rejection Antigens Presented By
CC       TITLE OF INVENTION:  HLA-B44 Molecules, And Uses Thereof
CC       NUMBER OF SEQUENCES:  30
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 3.5 inch, 360 kb storage
CC         COMPUTER:  IBM
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/796,883
CC         FILING DATE:  06-FEB-1997
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/602,506
CC         FILING DATE:  20-FEBRUARY-1996
CC         APPLICATION NUMBER:  08/531,864
CC         FILING DATE:  21-SEPTEMBER-1995
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/373,636
CC         FILING DATE:  17-JANUARY-1995
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  08/253,503
CC         FILING DATE:  3-JUNE-1994
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hanson, No. 5744353man D.
CC         REGISTRATION NUMBER:  30,946
CC         REFERENCE/DOCKET NUMBER:  LUD 5436
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  25:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  10 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
CC       FEATURE:
CC         NAME/KEY:  MAGE-1/HLA-B44
SQ   SEQUENCE   10 AA; 1104 MW; 684 CN;

  Query Match          100.0%;  Score 61;  DB 1;  Length 10;
  Best Local Similarity 100.0%;  Pred. No. 1.24e-01;
  Matches     9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      2 EADPTGHSY 10
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   19
ID   US-08-498-461-5     STANDARD;     PRT;   10 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 5, Application US/08498461
XX
CC    Sequence 5, Application US/08498461
CC    Patent No. 5827073
CC     GENERAL INFORMATION:
CC       APPLICANT:  Luescher, Immanuel; Anjuere, Fabienee;
CC       APPLICANT:  Layer, Andreas; Romero, Pedro; Cerottini, Jean-Charles
CC       TITLE OF INVENTION:  Photoreactive Peptide Derivatives
CC       NUMBER OF SEQUENCES:  16
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 3.5 inch, 1.44 kb storage
CC         COMPUTER:  IBM
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/498,461
CC         FILING DATE:  5-JULY-1995
CC         CLASSIFICATION:  435
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hanson, No. 5827073man D.
CC         REGISTRATION NUMBER:  30,946
CC         REFERENCE/DOCKET NUMBER:  LUD 5403
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  5:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  10 amino acids
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       FEATURE:
CC         OTHER INFORMATION:  The Xaa is iodinated 2,
CC         OTHER INFORMATION:  3-[4-azidosalicyloyl]-diaminopropionic acid.
SQ   SEQUENCE   10 AA; 1086 MW; 695 CN;

  Query Match          100.0%;  Score 61;  DB 2;  Length 10;
  Best Local Similarity 100.0%;  Pred. No. 1.24e-01;
  Matches     9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      2 EADPTGHSY 10
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   20
ID   US-08-560-024-4     STANDARD;     PRT;   12 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 4, Application US/08560024
XX
CC    Sequence 4, Application US/08560024
CC    Patent No. 5843448
CC     GENERAL INFORMATION:
CC       APPLICANT:  Chen, Yao-Tseng; Stockert, Elisabeth;
CC       APPLICANT:  Chen, Yachi; Garin-Chesa, Pilar; Rettig, Wolfgang J.;
CC       APPLICANT:  van der Bruggen, Pierre;  Boon-Falleur, Thierry;
CC       APPLICANT:  Old, Lloyd J.
CC       TITLE OF INVENTION:  MONOCLONAL ANTIBODIES WHICH BIND TO
CC       TITLE OF INVENTION:  TUMOR REJECTION ANTIGEN PRECURSOR MAGE-1, RECOMBINANT M
CC       TITLE OF INVENTION:  AND MAGE-1 DERIVED IMMUNOGENIC PEPTIDES
CC       NUMBER OF SEQUENCES:  4
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
CC         COMPUTER:  IBM
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/560,024
CC         FILING DATE:
```

```
CC         CLASSIFICATION:  514
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/190,411
CC         FILING DATE:  01-FEBRUARY-1994
CC         APPLICATION NUMBER:  037,230
CC         FILING DATE:  26-MARCH-1993
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US92/04354
CC         FILING DATE:  22-MAY-1992
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/807,043
CC         FILING DATE:  12-DECEMBER-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/764,364
CC         FILING DATE:  23-SEPTEMBER-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/728,838
CC         APPLICATION NUMBER:  9-JULY-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/705,702
CC         FILING DATE:  23-MAY-1991
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hanson, No. 5843448man D.
CC         REGISTRATION NUMBER:  30,946
CC         REFERENCE/DOCKET NUMBER:  LUD 5354
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  4:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  12 amino acid residues
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   12 AA; 1318 MW; 944 CN;

  Query Match          100.0%;  Score 61;  DB 2;  Length 12;
  Best Local Similarity 100.0%;  Pred. No. 1.24e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      4 EADPTGHSY 12
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   21
ID   US-08-190-411A-4     STANDARD;     PRT;   12 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 4, Application US/08190411A
XX
CC    Sequence 4, Application US/08190411A
CC    Patent No. 5541104
CC     GENERAL INFORMATION:
CC       APPLICANT:  Chen, Yao-Tseng; Stockert, Elisabeth;
CC       APPLICANT:  Chen, Yachi; Garin-Chesa, Pilar; Rettig, Wolfgang J.;
CC       APPLICANT:  van der Bruggen, Pierre;  Boon-Falleur, Thierry;
CC       APPLICANT:  Old, Lloyd J.
CC       TITLE OF INVENTION:  MONOCLONAL ANTIBODIES WHICH BIND TO
CC       TITLE OF INVENTION:  TUMOR REJECTION ANTIGEN PRECURSOR MAGE-1, RECOMBINANT MAGE
CC       TITLE OF INVENTION:  AND MAGE-1 DERIVED IMMUNOGENIC PEPTIDES
CC       NUMBER OF SEQUENCES:  4
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
CC         COMPUTER:  IBM
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/190,411A
CC         FILING DATE:  01-FEBRUARY-1994
CC         CLASSIFICATION:  436
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  037,230
CC         FILING DATE:  26-MARCH-1993
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  PCT/US92/04354
CC         FILING DATE:  22-MAY-1992
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/807,043
CC         FILING DATE:  12-DECEMBER-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/764,364
CC         FILING DATE:  23-SEPTEMBER-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/728,838
CC         APPLICATION NUMBER:  9-JULY-1991
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  07/705,702
CC         FILING DATE:  23-MAY-1991
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Hanson, No. 5541104man D.
CC         REGISTRATION NUMBER:  30,946
CC         REFERENCE/DOCKET NUMBER:  LUD 5354
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  4:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  12 amino acid residues
CC         TYPE:  amino acid
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   12 AA; 1318 MW; 944 CN;

  Query Match          100.0%;  Score 61;  DB 1;  Length 12;
  Best Local Similarity 100.0%;  Pred. No. 1.24e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      4 EADPTGHSY 12
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   22
ID   US-08-465-167A-24     STANDARD;     PRT;   309 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 24, Application US/08465167A
XX
CC    Sequence 24, Application US/08465167A
CC    Patent No. 5750395
CC     GENERAL INFORMATION:
CC       APPLICANT:  Fikes, John D.
CC       APPLICANT:  Livingston, Brian D.
CC       APPLICANT:  Sette, Alessandro D.
CC       APPLICANT:  Sidney, John C.
CC       TITLE OF INVENTION:  DNA ENCODING MAGE-1 C-TERMINAL
CC       TITLE OF INVENTION:  IMMUNOGENIC PEPTIDES (as amended)
CC       NUMBER OF SEQUENCES:  51
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Townsend and Townsend and Crew LLP
CC         STREET:  Two Embarcadero Center, 8th Floor
CC         CITY:  San Francisco
CC         STATE:  CA
```

```
CC         COUNTRY:  USA
CC         ZIP:  98111
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Floppy disk
CC         COMPUTER:  IBM PC compatible
CC         OPERATING SYSTEM:  PC-DOS/MS-DOS
CC         SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/465,167A
CC         FILING DATE:  05-JUN-1995
CC         CLASSIFICATION:  435
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US 08/103,623
CC         FILING DATE:  06-AUG-1993
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Parmelee, Steven W.
CC         REGISTRATION NUMBER:  31,990
CC         REFERENCE/DOCKET NUMBER:  14137-60-1
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  206-467-9600
CC         TELEFAX:  415-576-0300
CC     INFORMATION FOR SEQ ID NO:  24:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  309 amino acids
CC         TYPE:  amino acid
CC         STRANDEDNESS:  single
CC         TOPOLOGY:  linear
CC       MOLECULE TYPE:  protein
SQ   SEQUENCE   309 AA; 34342 MW; 512752 CN;

  Query Match          100.0%;  Score 61;  DB 1;  Length 309;
  Best Local Similarity 100.0%;  Pred. No. 1.24e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db    161 EADPTGHSY 169
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   23
ID   US-08-993-118-10     STANDARD;     PRT;   309 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 10, Application US/08993118
XX
CC    Sequence 10, Application US/08993118
CC    Patent No. 5997872
CC     GENERAL INFORMATION:
CC       APPLICANT:  LUCAS, Sophie;
CC       APPLICANT:  DE SMET, Charles;
CC       APPLICANT:  BOON-FALLEUR, Thierry
CC       TITLE OF INVENTION:  ISOLATED NUCLEIC ACID MOLECULE CODING FOR TUMOR
CC       TITLE OF INVENTION:  REJECTION ANTIGEN PRECURSOR MAGE-C1 AND USES
CC       TITLE OF INVENTION:  THEREOF
CC       NUMBER OF SEQUENCES:  14
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         COUNTRY:  USA
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 3.5 inch, 360 kb storage
CC         COMPUTER:  IBM PS/2
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/993,118
CC         FILING DATE:
CC         CLASSIFICATION:
CC       PRIOR APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/845,528
CC         FILING DATE:  April 25, 1997
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Mary Anne Schofield
CC         REGISTRATION NUMBER:  36,669
CC         REFERENCE/DOCKET NUMBER:  LUD 5455
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  10:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  309
CC         TYPE:  amino acids
CC         STRANDEDNESS:  single stranded
CC         TOPOLOGY:  linear
SQ   SEQUENCE   309 AA; 34342 MW; 512752 CN;

  Query Match          100.0%;  Score 61;  DB 2;  Length 309;
  Best Local Similarity 100.0%;  Pred. No. 1.24e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db    161 EADPTGHSY 169
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   24
ID   US-08-845-528C-10     STANDARD;     PRT;   309 AA.
XX
AC   xxxxxx
XX
DT
XX
DE   Sequence 10, Application US/08845528C
XX
CC    Sequence 10, Application US/08845528C
CC    Patent No. 6027924
CC     GENERAL INFORMATION:
CC       APPLICANT:  LUCAS, Sophie;
CC       APPLICANT:  DE SMET, Charles;
CC       APPLICANT:  BOON-FALLEUR, Thierry
CC       TITLE OF INVENTION:  ISOLATED NUCLEIC ACID MOLECULE CODING FOR TUMOR
CC       TITLE OF INVENTION:  REJECTION ANTIGEN PRECURSOR MAGE-C1 AND USES
CC       TITLE OF INVENTION:  THEREOF
CC       NUMBER OF SEQUENCES:  14
CC       CORRESPONDENCE ADDRESS:
CC         ADDRESSEE:  Felfe & Lynch
CC         STREET:  805 Third Avenue
CC         CITY:  New York City
CC         STATE:  New York
CC         COUNTRY:  USA
CC         ZIP:  10022
CC       COMPUTER READABLE FORM:
CC         MEDIUM TYPE:  Diskette, 3.5 inch, 360 kb storage
CC         COMPUTER:  IBM PS/2
CC         OPERATING SYSTEM:  PC-DOS
CC         SOFTWARE:  Wordperfect
CC       CURRENT APPLICATION DATA:
CC         APPLICATION NUMBER:  US/08/845,528C
CC         FILING DATE:  April 25, 1997
CC         CLASSIFICATION:  4335
CC       ATTORNEY/AGENT INFORMATION:
CC         NAME:  Mary Anne Schofield
CC         REGISTRATION NUMBER:  36,669
CC         REFERENCE/DOCKET NUMBER:  LUD 5455
CC       TELECOMMUNICATION INFORMATION:
CC         TELEPHONE:  (212) 688-9200
CC         TELEFAX:  (212) 838-3884
CC     INFORMATION FOR SEQ ID NO:  10:
CC       SEQUENCE CHARACTERISTICS:
CC         LENGTH:  309
```

```
CC         TYPE:  amino acids
CC         STRANDEDNESS:  single stranded
CC         TOPOLOGY:  linear
SQ   SEQUENCE   309 AA; 34342 MW; 512752 CN;

  Query Match          100.0%;  Score 61;  DB 3;  Length 309;
  Best Local Similarity 100.0%;  Pred. No. 1.24e-01;
  Matches     9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db    161 EADPTGHSY 169
          |||||||||
Qy      1 EADPTGHSY 9


Search completed: Wed Sep 13 07:14:50 2000
Job time : 8 secs.
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: September 13, 2000, 01:41:38 ; Search time 6530.16 Seconds
(without alignments)
1550.707 Million cell updates/sec

Title: US-08-819-669E-8
Perfect score: 5674
Sequence: 1 CCCGGGGCACCACTGGCATC.....TAATGATCTTGGGTGGATCC 5674

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 972840 seqs, 892348106 residues

Total number of hits satisfying chosen parameters: 1945680

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pr1:*
10: gb_pr2:*
11: gb_pr3:*
12: gb_ro:*
13: gb_sts:*
14: gb_sy:*
15: gb_un:*
16: em_fun:*
17: em_hum1:*
18: em_hum2:*
19: em_in:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_sy:*
29: em_un:*
30: em_vi:*
31: gb_htg1:*
32: gb_htg2:*
33: gb_in1:*
34: gb_in2:*
35: em_ba1:*
36: em_ba2:*
37: em_hum3:*
38: em_hum4:*
39: gb_pr4:*
40: gb_htg3:*
41: gb_htg4:*
42: gb_htg5:*
43: gb_htg6:*
44: gb_htg7:*
45: em_htg1:*
46: em_htg2:*
47: em_htg3:*
48: em_hum5:*
49: gb_pl3:*
50: gb_pr5:*
51: gb_htg8:*
52: gb_htg9:*
53: gb_htg10:*
54: gb_htg11:*
55: gb_htg12:*
56: gb_htg13:*
57: gb_htg14:*
58: gb_in3:*
59: gb_htg15:*
60: gb_htg16:*
61: gb_htg17:*
62: em_htg4:*
63: em_htg5:*
64: em_htg6:*
65: em_htg7:*
66: em_hum6:*
67: gb_htg18:*
68: gb_htg19:*
69: gb_htg20:*
70: gb_htg21:*
71: gb_htg22:*
72: gb_htg23:*
73: gb_htg24:*
74: gb_htg25:*
75: gb_htg26:*
76: gb_htg27:*
77: gb_htg28:*
78: gb_htg29:*
79: gb_htg30:*
80: gb_htg31:*
81: gb_vi1:*
82: gb_vi2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		5674	100.0	5674	5	AR060975	AR060975 Sequence
2		5674	100.0	5674	5	I24013	I24013 Sequence 1
3		5674	100.0	5674	5	I36923	I36923 Sequence 8
4	c	5566.2	98.1	275159	50	U82670	U82670 Homo sapien
5	c	5532.6	97.5	156854	48	HSU82672	U82672 Human chrom
6		2777.2	48.9	49375	39	AF134576	AF134576 Homo sapi
7		2655	46.8	11495	10	HSU10687	U10687 Human MAGE-
8		2653.2	46.8	158898	11	U82696	U82696 Homo sapien
9		2513.6	44.3	4895	10	HSU10688	U10688 Human MAGE-
10		2429.6	42.8	4736	10	HSU10690	U10690 Human MAGE-
11		2422.8	42.7	4741	10	HSU10689	U10689 Human MAGE-
12		2419	42.6	2419	5	I36922	I36922 Sequence 7
13		2408	42.4	2420	5	AR007331	AR007331 Sequence
14		2408	42.4	2420	9	HUMMAG1A	M77481 Human antig
15		2328	41.0	178515	74	AC009621	AC009621 Homo sapi
16	c	2307.6	40.7	37781	17	AF030261	Af030261 Homo sapi
17		2307.6	40.7	54155	37	HSAF2994	Af002994 Homo sapi
18	c	2307.6	40.7	302250	50	U82671	U82671 Homo sapien
19		2306	40.6	48574	37	HSAF2996	Af002996 Homo sapi
20		2306	40.6	302250	50	U82671	U82671 Homo sapien
21	c	2199.6	38.8	48574	37	HSAF2996	Af002996 Homo sapi
22		2117.8	37.3	4559	9	HUMMAGE2X	L18920 Human MAGE-
23		2058.8	36.3	4157	5	I36924	I36924 Sequence 9
24		1945	34.3	4523	9	HUMMAGE12X	L18877 Human MAGE-

25	1762.6	31.1	4204	9	HSU03735	U03735 Human MAGE-
26	1678.8	29.6	2531	5	I36928	I36928 Sequence 13
27	1678.8	29.6	2531	5	I36929	I36929 Sequence 14
28	1677	29.6	3871	10	HSU10691	U10691 Human MAGE-
29	1556.6	27.4	43927	11	HSU69568	U69568 Human Xq28
30	1555	27.4	118440	50	AC016939	AC016939 Homo sapi
31	1536.4	27.1	3680	10	HSU10692	U10692 Human MAGE-
32	1535.2	27.1	111560	50	AC016940	AC016940 Homo sapi
33	1504.6	26.5	2305	5	I36932	I36932 Sequence 17
34	1410.8	24.9	2226	5	I36931	I36931 Sequence 16
35	1396.6	24.6	74299	50	AC016941	AC016941 Homo sapi
36	1353.8	23.9	73360	11	HSU66083	U66083 Human conti
37	1247.4	22.0	3839	10	HSU10693	U10693 Human MAGE-
38	1226.6	21.6	2931	10	HSU10694	U10694 Human MAGE-
39	1051	18.5	3672	10	HSU10686	U10686 Human MAGE-
40	988.8	17.4	1640	5	I36926	I36926 Sequence 11
c 41	955.8	16.8	165648	69	AC024727	AC024727 Homo sapi
42	900.2	15.9	1810	5	I36935	I36935 Sequence 20
43	866	15.3	3510	10	HSU10685	U10685 Human MAGE-
44	863.4	15.2	165648	69	AC024727	AC024727 Homo sapi
45	834.4	14.7	1094	5	AR052773	AR052773 Sequence

ALIGNMENTS

```
RESULT   1
AR060975
LOCUS       AR060975     5674 bp    DNA             PAT       29-SEP-1999
DEFINITION  Sequence 1 from patent US 5843448.
ACCESSION   AR060975
VERSION     AR060975.1  GI:5988666
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 5674)
  AUTHORS   Chen,Y., Stockert,E., Chen,Y., Garin-Chesa,P., Rettig,W.J. and
            Old,L.J.
  TITLE     Tumor rejection antigen precursor
  JOURNAL   Patent: US 5843448-A 1 01-DEC-1998;
FEATURES             Location/Qualifiers
     source          1. .5674
                     /organism="unknown"
BASE COUNT     1276 a   1644 c   1569 g   1185 t
ORIGIN
```

Query Match 100.0%; Score 5674; DB 5; Length 5674;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 CCCGGGGCACCACTGGCATCCCTCCCCCTACCACCCCCAATCCCTCCCTTTACGCCACCC 60
Db       1 CCCGGGGCACCACTGGCATCCCTCCCCCTACCACCCCCAATCCCTCCCTTTACGCCACCC 60

Qy      61 ATCCAAACATCTTCACGCTCACCCCCAGCCCAAGCCAGGCAGAATCCGGTTCCACCCCTG 120
Db      61 ATCCAAACATCTTCACGCTCACCCCCAGCCCAAGCCAGGCAGAATCCGGTTCCACCCCTG 120

Qy     121 CTCTCAACCCAGGGAAGCCCAGGTGCCCAGATGTGACGCCACTGACTTGAGCATTAGTGG 180
Db     121 CTCTCAACCCAGGGAAGCCCAGGTGCCCAGATGTGACGCCACTGACTTGAGCATTAGTGG 180

Qy     181 TTAGAGAGAAGCGAGGTTTTCGGTCTGAGGGGCGGCTTGAGATCGGTGGAGGGAAGCGGG 240
Db     181 TTAGAGAGAAGCGAGGTTTTCGGTCTGAGGGGCGGCTTGAGATCGGTGGAGGGAAGCGGG 240

Qy     241 CCCAGCTCTGTAAGGAGGCAAGGTGACATGCTGAGGGAGGACTGAGGACCCACTTACCCC 300
Db     241 CCCAGCTCTGTAAGGAGGCAAGGTGACATGCTGAGGGAGGACTGAGGACCCACTTACCCC 300

Qy     301 AGATAGAGGACCCCAAATAATCCCTTCATGCCAGTCCTGGACCATCTGGTGGTGGACTTC 360
Db     301 AGATAGAGGACCCCAAATAATCCCTTCATGCCAGTCCTGGACCATCTGGTGGTGGACTTC 360

Qy     361 TCAGGCTGGGCCACCCCCAGCCCCCTTGCTGCTTAAACCACTGGGGACTCGAAGTCAGAG 420
Db     361 TCAGGCTGGGCCACCCCCAGCCCCCTTGCTGCTTAAACCACTGGGGACTCGAAGTCAGAG 420

Qy     421 CTCCGTGTGATCAGGGAAGGGCTGCTTAGGAGAGGGCAGCGTCCAGGCTCTGCCAGACAT 480
Db     421 CTCCGTGTGATCAGGGAAGGGCTGCTTAGGAGAGGGCAGCGTCCAGGCTCTGCCAGACAT 480

Qy     481 CATGCTCAGGATTCTCAAGGAGGGCTGAGGGTCCCTAAGACCCCACTCCCGTGACCCAAC 540
Db     481 CATGCTCAGGATTCTCAAGGAGGGCTGAGGGTCCCTAAGACCCCACTCCCGTGACCCAAC 540

Qy     541 CCCCACTCCAATGCTCACTCCCGTGACCCAACCCCCTCTTCATTGTCATTCCAACCCCCA 600
Db     541 CCCCACTCCAATGCTCACTCCCGTGACCCAACCCCCTCTTCATTGTCATTCCAACCCCCA 600

Qy     601 CCCCACATCCCCCACCCCATCCCTCAACCCTGATGCCCATCCGCCCAGCCATTCCACCCT 660
Db     601 CCCCACATCCCCCACCCCATCCCTCAACCCTGATGCCCATCCGCCCAGCCATTCCACCCT 660

Qy     661 CACCCCCACCCCCACCCCCACGCCCACTCCCACCCCCACCCAGGCAGGATCCGGTTCCCG 720
Db     661 CACCCCCACCCCCACCCCCACGCCCACTCCCACCCCCACCCAGGCAGGATCCGGTTCCCG 720

Qy     721 CCAGGAAACATCCGGGTGCCCGGATGTGACGCCACTGACTTGCGCATTGTGGGGCAGAGA 780
Db     721 CCAGGAAACATCCGGGTGCCCGGATGTGACGCCACTGACTTGCGCATTGTGGGGCAGAGA 780

Qy     781 GAAGCGAGGTTTCCATTCTGAGGGACGGCGTAGAGTTCGGCCGAAGGAACCTGACCCAGG 840
Db     781 GAAGCGAGGTTTCCATTCTGAGGGACGGCGTAGAGTTCGGCCGAAGGAACCTGACCCAGG 840

Qy     841 CTCTGTGAGGAGGCAAGGTGAGAGGCTGAGGGAGGACTGAGGACCCCGCCACTCCAAATA 900
Db     841 CTCTGTGAGGAGGCAAGGTGAGAGGCTGAGGGAGGACTGAGGACCCCGCCACTCCAAATA 900

Qy     901 GAGAGCCCCAAATATTCCAGCCCCGCCCTTGCTGCCAGCCCTGGCCCACCCGCGGGAAGA 960
Db     901 GAGAGCCCCAAATATTCCAGCCCCGCCCTTGCTGCCAGCCCTGGCCCACCCGCGGGAAGA 960

Qy     961 CGTCTCAGCCTGGGCTGCCCCCAGACCCCTGCTCCAAAAGCCTTGAGAGACACCAGGTTC 1020
Db     961 CGTCTCAGCCTGGGCTGCCCCCAGACCCCTGCTCCAAAAGCCTTGAGAGACACCAGGTTC 1020

Qy    1021 TTCTCCCCAAGCTCTGGAATCAGAGGTTGCTGTGACCAGGGCAGGACTGGTTAGGAGAGG 1080
Db    1021 TTCTCCCCAAGCTCTGGAATCAGAGGTTGCTGTGACCAGGGCAGGACTGGTTAGGAGAGG 1080

Qy    1081 GCAGGGCACAGGCTCTGCCAGGCATCAAGATCAGCACCCAAGAGGGAGGGCTGTGGGCCC 1140
Db    1081 GCAGGGCACAGGCTCTGCCAGGCATCAAGATCAGCACCCAAGAGGGAGGGCTGTGGGCCC 1140

Qy    1141 CCAAGACTGCACTCCAATCCCCACTCCCACCCCATTCGCATTCCCATTCCCCACCCAACC 1200
Db    1141 CCAAGACTGCACTCCAATCCCCACTCCCACCCCATTCGCATTCCCATTCCCCACCCAACC 1200

Qy    1201 CCCATCTCCTCAGCTACACCTCCACCCCCATCCCTACTCCTACTCCGTCACCTGACCACC 1260
Db    1201 CCCATCTCCTCAGCTACACCTCCACCCCCATCCCTACTCCTACTCCGTCACCTGACCACC 1260

Qy    1261 ACCCTCCAGCCCCAGCACCAGCCCCAACCCTTCTGCCACCTCACCCTCACTGCCCCCAAC 1320
Db    1261 ACCCTCCAGCCCCAGCACCAGCCCCAACCCTTCTGCCACCTCACCCTCACTGCCCCCAAC 1320

Qy    1321 CCCACCCTCATCTCTCTCATGTGCCCCACTCCCATCGCCTCCCCCATTCTGGCAGAATCC 1380
Db    1321 CCCACCCTCATCTCTCTCATGTGCCCCACTCCCATCGCCTCCCCCATTCTGGCAGAATCC 1380

Qy    1381 GGTTTGCCCCTGCTCTCAACCCAGGGAAGCCCTGGTAGGCCCGATGTGAAACCACTGACT 1440
```

```
Db   1381 GGTTTGCCCCTGCTCTCAACCCAGGGAAGCCCTGGTAGGCCCGATGTGAAACCACTGACT 1440
Qy   1441 TGAACCTCACAGATCTGAGAGAAGCCAGGTTCATTTAATGGTTCTGAGGGGCGGCTTGAG 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 TGAACCTCACAGATCTGAGAGAAGCCAGGTTCATTTAATGGTTCTGAGGGGCGGCTTGAG 1500
Qy   1501 ATCCACTGAGGGGAGTGGTTTTAGGCTCTGTGAGGAGGCAAGGTGAGATGCTGAGGGAGG 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 ATCCACTGAGGGGAGTGGTTTTAGGCTCTGTGAGGAGGCAAGGTGAGATGCTGAGGGAGG 1560
Qy   1561 ACTGAGGAGGCACACACCCCAGGTAGATGGCCCCAAAATGATCCAGTACCACCCCTGCTG 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 ACTGAGGAGGCACACACCCCAGGTAGATGGCCCCAAAATGATCCAGTACCACCCCTGCTG 1620
Qy   1621 CCAGCCCTGGACCACCCGGCCAGGACAGATGTCTCAGCTGGACCACCCCCCGTCCCGTCC 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 CCAGCCCTGGACCACCCGGCCAGGACAGATGTCTCAGCTGGACCACCCCCCGTCCCGTCC 1680
Qy   1681 CACTGCCACTTAACCCACAGGGCAATCTGTAGTCATAGCTTATGTGACCGGGGCAGGGTT 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 CACTGCCACTTAACCCACAGGGCAATCTGTAGTCATAGCTTATGTGACCGGGGCAGGGTT 1740
Qy   1741 GGTCAGGAGAGGCAGGGCCCAGGCATCAAGGTCCAGCATCCGCCCGGCATTAGGGTCAGG 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 GGTCAGGAGAGGCAGGGCCCAGGCATCAAGGTCCAGCATCCGCCCGGCATTAGGGTCAGG 1800
Qy   1801 ACCCTGGGAGGGAACTGAGGGTTCCCCACCCACACCTGTCTCCTCATCTCCACCGCCACC 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 ACCCTGGGAGGGAACTGAGGGTTCCCCACCCACACCTGTCTCCTCATCTCCACCGCCACC 1860
Qy   1861 CCACTCACATTCCCATACCTACCCCCTACCCCCAACCTCATCTTGTCAGAATCCCTGCTG 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 CCACTCACATTCCCATACCTACCCCCTACCCCCAACCTCATCTTGTCAGAATCCCTGCTG 1920
Qy   1921 TCAACCCACGGAAGCCACGGGAATGGCGGCCAGGCACTCGGATCTTGACGTCCCCATCCA 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 TCAACCCACGGAAGCCACGGGAATGGCGGCCAGGCACTCGGATCTTGACGTCCCCATCCA 1980
Qy   1981 GGGTCTGATGGAGGGAAGGGGCTTGAACAGGGCCTCAGGGGAGCAGAGGGAGGGCCCTAC 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 GGGTCTGATGGAGGGAAGGGGCTTGAACAGGGCCTCAGGGGAGCAGAGGGAGGGCCCTAC 2040
Qy   2041 TGCGAGATGAGGGAGGCCTCAGAGGACCCAGCACCCTAGGACACCGCACCCCTGTCTGAG 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 TGCGAGATGAGGGAGGCCTCAGAGGACCCAGCACCCTAGGACACCGCACCCCTGTCTGAG 2100
Qy   2101 ACTGAGGCTGCCACTTCTGGCCTCAAGAATCAGAACGATGGGGACTCAGATTGCATGGGG 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 ACTGAGGCTGCCACTTCTGGCCTCAAGAATCAGAACGATGGGGACTCAGATTGCATGGGG 2160
Qy   2161 GTGGGACCCAGGCCTGCAAGGCTTACGCGGAGGAAGAGGAGGGAGGACTCAGGGGACCTT 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 GTGGGACCCAGGCCTGCAAGGCTTACGCGGAGGAAGAGGAGGGAGGACTCAGGGGACCTT 2220
Qy   2221 GGAATCCAGATCAGTGTGGACCTCGGCCCTGAGAGGTCCAGGGCACGGTGGCCACATATG 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 GGAATCCAGATCAGTGTGGACCTCGGCCCTGAGAGGTCCAGGGCACGGTGGCCACATATG 2280
Qy   2281 GCCCATATTTCCTGCATCTTTGAGGTGACAGGACAGAGCTGTGGTCTGAGAAGTGGGGCC 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 GCCCATATTTCCTGCATCTTTGAGGTGACAGGACAGAGCTGTGGTCTGAGAAGTGGGGCC 2340
Qy   2341 TCAGGTCAACAGAGGGAGGAGTTCCAGGATCCATATGGCCCAAGATGTGCCCCCTTCATG 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 TCAGGTCAACAGAGGGAGGAGTTCCAGGATCCATATGGCCCAAGATGTGCCCCCTTCATG 2400
Qy   2401 AGGACTGGGGATATCCCCGGCTCAGAAAGAAGGGACTCCACACAGTCTGGCTGTCCCCTT 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 AGGACTGGGGATATCCCCGGCTCAGAAAGAAGGGACTCCACACAGTCTGGCTGTCCCCTT 2460
Qy   2461 TTAGTAGCTCTAGGGGGACCAGATCAGGGATGGCGGTATGTTCCATTCTCACTTGTACCA 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 TTAGTAGCTCTAGGGGGACCAGATCAGGGATGGCGGTATGTTCCATTCTCACTTGTACCA 2520

Qy   2521 CAGGCAGGAAGTTGGGGGGCCCTCAGGGAGATGGGGTCTTGGGGTAAAGGGGGGATGTCT 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 CAGGCAGGAAGTTGGGGGGCCCTCAGGGAGATGGGGTCTTGGGGTAAAGGGGGGATGTCT 2580
Qy   2581 ACTCATGTCAGGGAATTGGGGGTTGAGGAAGCACAGGCGCTGGCAGGAATAAAGATGAGT 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 ACTCATGTCAGGGAATTGGGGGTTGAGGAAGCACAGGCGCTGGCAGGAATAAAGATGAGT 2640
Qy   2641 GAGACAGACAAGGCTATTGGAATCCACACCCCAGAACCAAAGGGGTCAGCCCTGGACACC 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 GAGACAGACAAGGCTATTGGAATCCACACCCCAGAACCAAAGGGGTCAGCCCTGGACACC 2700
Qy   2701 TCACCCAGGATGTGGCTTCTTTTTCACTCCTGTTTCCAGATCTGGGGCAGGTGAGGACCT 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 TCACCCAGGATGTGGCTTCTTTTTCACTCCTGTTTCCAGATCTGGGGCAGGTGAGGACCT 2760
Qy   2761 CATTCTCAGAGGGTGACTCAGGTCAACGTAGGGACCCCCATCTGGTCTAAAGACAGAGCG 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 CATTCTCAGAGGGTGACTCAGGTCAACGTAGGGACCCCCATCTGGTCTAAAGACAGAGCG 2820
Qy   2821 GTCCCAGGATCTGCCATGCGTTCGGGTGAGGAACATGAGGGAGGACTGAGGGTACCCCAG 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 GTCCCAGGATCTGCCATGCGTTCGGGTGAGGAACATGAGGGAGGACTGAGGGTACCCCAG 2880
Qy   2881 GACCAGAACACTGAGGGAGACTGCACAGAAATCAGCCCTGCCCCTGCTGTCACCCCAGAG 2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881 GACCAGAACACTGAGGGAGACTGCACAGAAATCAGCCCTGCCCCTGCTGTCACCCCAGAG 2940
Qy   2941 AGCATGGGCTGGGCCGTCTGCCGAGGTCCTTCCGTTATCCTGGGATCATTGATGTCAGGG 3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941 AGCATGGGCTGGGCCGTCTGCCGAGGTCCTTCCGTTATCCTGGGATCATTGATGTCAGGG 3000
Qy   3001 ACGGGGAGGCCTTGGTCTGAGAAGGCTGCGCTCAGGTCAGTAGAGGGAGCGTCCCAGGCC 3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3001 ACGGGGAGGCCTTGGTCTGAGAAGGCTGCGCTCAGGTCAGTAGAGGGAGCGTCCCAGGCC 3060
Qy   3061 CTGCCAGGAGTCAAGGTGAGGACCAAGCGGGCACCTCACCCAGGACACATTAATTCCAAT 3120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3061 CTGCCAGGAGTCAAGGTGAGGACCAAGCGGGCACCTCACCCAGGACACATTAATTCCAAT 3120
Qy   3121 GAATTTTGATATCTCTTGCTGCCCTTCCCCAAGGACCTAGGCACGTGTGGCCAGATGTTT 3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3121 GAATTTTGATATCTCTTGCTGCCCTTCCCCAAGGACCTAGGCACGTGTGGCCAGATGTTT 3180
Qy   3181 GTCCCCTCCTGTCCTTCCATTCCTTATCATGGATGTGAACTCTTGATTTGGATTTCTCAG 3240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3181 GTCCCCTCCTGTCCTTCCATTCCTTATCATGGATGTGAACTCTTGATTTGGATTTCTCAG 3240
Qy   3241 ACCAGCAAAAGGGCAGGATCCAGGCCCTGCCAGGAAAAATATAAGGGCCCTGCGTGAGAA 3300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3241 ACCAGCAAAAGGGCAGGATCCAGGCCCTGCCAGGAAAAATATAAGGGCCCTGCGTGAGAA 3300
Qy   3301 CAGAGGGGGTCATCCACTGCATGAGAGTGGGGATGTCACAGAGTCCAGCCCACCCTCCTG 3360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3301 CAGAGGGGGTCATCCACTGCATGAGAGTGGGGATGTCACAGAGTCCAGCCCACCCTCCTG 3360
Qy   3361 GTAGCACTGAGAAGCCAGGGCTGTGCTTGCGGTCTGCACCCTGAGGGCCCGTGGATTCCT 3420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3361 GTAGCACTGAGAAGCCAGGGCTGTGCTTGCGGTCTGCACCCTGAGGGCCCGTGGATTCCT 3420
Qy   3421 CTTCCTGGAGCTCCAGGAACCAGGCAGTGAGGCCTTGGTCTGAGACAGTATCCTCAGGTC 3480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3421 CTTCCTGGAGCTCCAGGAACCAGGCAGTGAGGCCTTGGTCTGAGACAGTATCCTCAGGTC 3480
Qy   3481 ACAGAGCAGAGGATGCACAGGGTGTGCCAGCAGTGAATGTTTGCCCTGAATGCACACCAA 3540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3481 ACAGAGCAGAGGATGCACAGGGTGTGCCAGCAGTGAATGTTTGCCCTGAATGCACACCAA 3540
Qy   3541 GGGCCCCACCTGCCACAGGACACATAGGACTCCACAGAGTCTGGCCTCACCTCCCTACTG 3600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3541 GGGCCCCACCTGCCACAGGACACATAGGACTCCACAGAGTCTGGCCTCACCTCCCTACTG 3600
```

```
Qy   3601 TCAGTCCTGTAGAATCGACCTCTGCTGGCCGGCTGTACCCTGAGTACCCTCTCACTTCCT 3660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3601 TCAGTCCTGTAGAATCGACCTCTGCTGGCCGGCTGTACCCTGAGTACCCTCTCACTTCCT 3660

Qy   3661 CCTTCAGGTTTTCAGGGGACAGGCCAACCCAGAGGACAGGATTCCCTGGAGGCCACAGAG 3720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3661 CCTTCAGGTTTTCAGGGGACAGGCCAACCCAGAGGACAGGATTCCCTGGAGGCCACAGAG 3720

Qy   3721 GAGCACCAAGGAGAAGATCTGTAAGTAGGCCTTTGTTAGAGTCTCCAAGGTTCAGTTCTC 3780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3721 GAGCACCAAGGAGAAGATCTGTAAGTAGGCCTTTGTTAGAGTCTCCAAGGTTCAGTTCTC 3780

Qy   3781 AGCTGAGGCCTCTCACACACTCCCTCTCTCCCCAGGCCTGTGGGTCTTCATTGCCCAGCT 3840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3781 AGCTGAGGCCTCTCACACACTCCCTCTCTCCCCAGGCCTGTGGGTCTTCATTGCCCAGCT 3840

Qy   3841 CCTGCCCACACTCCTGCCTGCTGCCCTGACGAGAGTCATCATGTCTCTTGAGCAGAGGAG 3900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3841 CCTGCCCACACTCCTGCCTGCTGCCCTGACGAGAGTCATCATGTCTCTTGAGCAGAGGAG 3900

Qy   3901 TCTGCACTGCAAGCCTGAGGAAGCCCTTGAGGCCCAACAAGAGGCCCTGGGCCTGGTGTG 3960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3901 TCTGCACTGCAAGCCTGAGGAAGCCCTTGAGGCCCAACAAGAGGCCCTGGGCCTGGTGTG 3960

Qy   3961 TGTGCAGGCTGCCACCTCCTCCTCCTCTCCTCTGGTCCTGGGCACCCTGGAGGAGGTGCC 4020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3961 TGTGCAGGCTGCCACCTCCTCCTCCTCTCCTCTGGTCCTGGGCACCCTGGAGGAGGTGCC 4020

Qy   4021 CACTGCTGGGTCAACAGATCCTCCCCAGAGTCCTCAGGGAGCCTCCGCCTTTCCCACTAC 4080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4021 CACTGCTGGGTCAACAGATCCTCCCCAGAGTCCTCAGGGAGCCTCCGCCTTTCCCACTAC 4080

Qy   4081 CATCAACTTCACTCGACAGAGGCAACCCAGTGAGGGTTCCAGCAGCCGTGAAGAGGAGGG 4140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4081 CATCAACTTCACTCGACAGAGGCAACCCAGTGAGGGTTCCAGCAGCCGTGAAGAGGAGGG 4140

Qy   4141 GCCAAGCACCTCTTGTATCCTGGAGTCCTTGTTCCGAGCAGTAATCACTAAGAAGGTGGC 4200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4141 GCCAAGCACCTCTTGTATCCTGGAGTCCTTGTTCCGAGCAGTAATCACTAAGAAGGTGGC 4200

Qy   4201 TGATTTGGTTGGTTTTCTGCTCCTCAAATATCGAGCCAGGGAGCCAGTCACAAAGGCAGA 4260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4201 TGATTTGGTTGGTTTTCTGCTCCTCAAATATCGAGCCAGGGAGCCAGTCACAAAGGCAGA 4260

Qy   4261 AATGCTGGAGAGTGTCATCAAAAATTACAAGCACTGTTTTCCTGAGATCTTCGGCAAAGC 4320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4261 AATGCTGGAGAGTGTCATCAAAAATTACAAGCACTGTTTTCCTGAGATCTTCGGCAAAGC 4320

Qy   4321 CTCTGAGTCCTTGCAGCTGGTCTTTGGCATTGACGTGAAGGAAGCAGACCCCACCGGCCA 4380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4321 CTCTGAGTCCTTGCAGCTGGTCTTTGGCATTGACGTGAAGGAAGCAGACCCCACCGGCCA 4380

Qy   4381 CTCCTATGTCCTTGTCACCTGCCTAGGTCTCTCCTATGATGGCCTGCTGGGTGATAATCA 4440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4381 CTCCTATGTCCTTGTCACCTGCCTAGGTCTCTCCTATGATGGCCTGCTGGGTGATAATCA 4440

Qy   4441 GATCATGCCCAAGACAGGCTTCCTGATAATTGTCCTGGTCATGATTGCAATGGAGGGCGG 4500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4441 GATCATGCCCAAGACAGGCTTCCTGATAATTGTCCTGGTCATGATTGCAATGGAGGGCGG 4500

Qy   4501 CCATGCTCCTGAGGAGGAAATCTGGGAGGAGCTGAGTGTGATGGAGGTGTATGATGGGAG 4560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4501 CCATGCTCCTGAGGAGGAAATCTGGGAGGAGCTGAGTGTGATGGAGGTGTATGATGGGAG 4560

Qy   4561 GGAGCACAGTGCCTATGGGGAGCCCAGGAAGCTGCTCACCCAAGATTTGGTGCAGGAAAA 4620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4561 GGAGCACAGTGCCTATGGGGAGCCCAGGAAGCTGCTCACCCAAGATTTGGTGCAGGAAAA 4620

Qy   4621 GTACCTGGAGTACGGCAGGTGCCGGACAGTGATCCCGCACGCTATGAGTTCCTGTGGGGT 4680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4621 GTACCTGGAGTACGGCAGGTGCCGGACAGTGATCCCGCACGCTATGAGTTCCTGTGGGGT 4680

Qy   4681 CCAAGGGCCCTCGCTGAAACCAGCTATGTGAAAGTCCTTGAGTATGTGATCAAGGTCAGT 4740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4681 CCAAGGGCCCTCGCTGAAACCAGCTATGTGAAAGTCCTTGAGTATGTGATCAAGGTCAGT 4740

Qy   4741 GCAAGAGTTCGCTTTTTCTTCCCATCCCTGCGTGAAGCAGCTTTGAGAGAGGAGGAAGAG 4800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4741 GCAAGAGTTCGCTTTTTCTTCCCATCCCTGCGTGAAGCAGCTTTGAGAGAGGAGGAAGAG 4800

Qy   4801 GGAGTCTGAGCATGAGTTGCAGCCAAGGCCAGTGGGAGGGGGACTGGGCCAGTGCACCTT 4860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4801 GGAGTCTGAGCATGAGTTGCAGCCAAGGCCAGTGGGAGGGGGACTGGGCCAGTGCACCTT 4860

Qy   4861 CCAGGGCCGCGTCCAGCAGCTTCCCCTGCCTCGTGTGACATGAGGCCCATTCTTCACTCT 4920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4861 CCAGGGCCGCGTCCAGCAGCTTCCCCTGCCTCGTGTGACATGAGGCCCATTCTTCACTCT 4920

Qy   4921 GAAGAGAGCGGTCAGTGTTCTCAGTAGTAGGTTTCTGTTCTATTGGGTGACTTGGAGATT 4980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4921 GAAGAGAGCGGTCAGTGTTCTCAGTAGTAGGTTTCTGTTCTATTGGGTGACTTGGAGATT 4980

Qy   4981 TATCTTTGTTCTCTTTTGGAATTGTTCAAATGTTTTTTTTTAAGGGATGGTTGAATGAAC 5040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4981 TATCTTTGTTCTCTTTTGGAATTGTTCAAATGTTTTTTTTTAAGGGATGGTTGAATGAAC 5040

Qy   5041 TTCAGCATCCAAGTTTATGAATGACAGCAGTCACACAGTTCTGTGTATATAGTTTAAGGG 5100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5041 TTCAGCATCCAAGTTTATGAATGACAGCAGTCACACAGTTCTGTGTATATAGTTTAAGGG 5100

Qy   5101 TAAGAGTCTTGTGTTTTATTCAGATTGGGAAATCCATTCTATTTTGTGAATTGGGATAAT 5160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5101 TAAGAGTCTTGTGTTTTATTCAGATTGGGAAATCCATTCTATTTTGTGAATTGGGATAAT 5160

Qy   5161 AACAGCAGTGGAATAAGTACTTAGAAATGTGAAAAATGAGCAGTAAAATAGATGAGATAA 5220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5161 AACAGCAGTGGAATAAGTACTTAGAAATGTGAAAAATGAGCAGTAAAATAGATGAGATAA 5220

Qy   5221 AGAACTAAAGAAATTAAGAGATAGTCAATTCTTGCCTTATACCTCAGTCTATTCTGTAAA 5280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5221 AGAACTAAAGAAATTAAGAGATAGTCAATTCTTGCCTTATACCTCAGTCTATTCTGTAAA 5280

Qy   5281 ATTTTTAAAGATATATGCATACCTGGATTTCCTTGGCTTCTTTGAGAATGTAAGAGAAAT 5340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5281 ATTTTTAAAGATATATGCATACCTGGATTTCCTTGGCTTCTTTGAGAATGTAAGAGAAAT 5340

Qy   5341 TAAATCTGAATAAAGAATTCTTCCTGTTCACTGGCTCTTTTCTTCTCCATGCACTGAGCA 5400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5341 TAAATCTGAATAAAGAATTCTTCCTGTTCACTGGCTCTTTTCTTCTCCATGCACTGAGCA 5400

Qy   5401 TCTGCTTTTTGGAAGGCCCTGGGTTAGTAGTGGAGATGCTAAGGTAAGCCAGACTCATAC 5460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5401 TCTGCTTTTTGGAAGGCCCTGGGTTAGTAGTGGAGATGCTAAGGTAAGCCAGACTCATAC 5460

Qy   5461 CCACCCATAGGGTCGTAGAGTCTAGGAGCTGCAGTCACGTAATCGAGGTGGCAAGATGTC 5520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5461 CCACCCATAGGGTCGTAGAGTCTAGGAGCTGCAGTCACGTAATCGAGGTGGCAAGATGTC 5520

Qy   5521 CTCTAAAGATGTAGGGAAAAGTGAGAGAGGGGTGAGGGTGTGGGGCTCCGGGTGAGAGTG 5580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5521 CTCTAAAGATGTAGGGAAAAGTGAGAGAGGGGTGAGGGTGTGGGGCTCCGGGTGAGAGTG 5580

Qy   5581 GTGGAGTGTCAATGCCCTGAGCTGGGGCATTTTGGGCTTTGGGAAACTGCAGTTCCTTCT 5640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5581 GTGGAGTGTCAATGCCCTGAGCTGGGGCATTTTGGGCTTTGGGAAACTGCAGTTCCTTCT 5640

Qy   5641 GGGGGAGCTGATTGTAATGATCTTGGGTGGATCC 5674
          |||||||||||||||||||||||||||||||||
Db   5641 GGGGGAGCTGATTGTAATGATCTTGGGTGGATCC 5674
```

RESULT 2
I24013
LOCUS I24013 5674 bp DNA PAT 07-OCT-1996
DEFINITION Sequence 1 from patent US 5541104.
ACCESSION I24013

```
VERSION     I24013.1  GI:1603883
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 5674)
  AUTHORS   Chen,Y., Stockert,E., Chen,Y., Garin-Chesa,P., Rettig,W.J., van der
            Bruggen,P., Boon-Falleur,T. and Old,L.J.
  TITLE     Monoclonal antibodies which bind to tumor rejection antigen
            precursor mage-1
  JOURNAL   Patent: US 5541104-A 1 30-JUL-1996;
FEATURES             Location/Qualifiers
     source          1. .5674
                     /organism="unknown"
BASE COUNT     1276 a   1644 c   1569 g   1185 t
ORIGIN
```

Query Match 100.0%; Score 5674; DB 5; Length 5674;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 CCCGGGGCACCACTGGCATCCCTCCCCCTACCACCCCCAATCCCTCCCTTTACGCCACCC 60
Db       1 CCCGGGGCACCACTGGCATCCCTCCCCCTACCACCCCCAATCCCTCCCTTTACGCCACCC 60

Qy      61 ATCCAAACATCTTCACGCTCACCCCCAGCCCAAGCCAGGCAGAATCCGGTTCCACCCCTG 120
Db      61 ATCCAAACATCTTCACGCTCACCCCCAGCCCAAGCCAGGCAGAATCCGGTTCCACCCCTG 120

Qy     121 CTCTCAACCCAGGGAAGCCCAGGTGCCCAGATGTGACGCCACTGACTTGAGCATTAGTGG 180
Db     121 CTCTCAACCCAGGGAAGCCCAGGTGCCCAGATGTGACGCCACTGACTTGAGCATTAGTGG 180

Qy     181 TTAGAGAGAAGCGAGGTTTTCGGTCTGAGGGGCGGCTTGAGATCGGTGGAGGGAAGCGGG 240
Db     181 TTAGAGAGAAGCGAGGTTTTCGGTCTGAGGGGCGGCTTGAGATCGGTGGAGGGAAGCGGG 240

Qy     241 CCCAGCTCTGTAAGGAGGCAAGGTGACATGCTGAGGGAGGACTGAGGACCCACTTACCCC 300
Db     241 CCCAGCTCTGTAAGGAGGCAAGGTGACATGCTGAGGGAGGACTGAGGACCCACTTACCCC 300

Qy     301 AGATAGAGGACCCCAAATAATCCCTTCATGCCAGTCCTGGACCATCTGGTGGTGGACTTC 360
Db     301 AGATAGAGGACCCCAAATAATCCCTTCATGCCAGTCCTGGACCATCTGGTGGTGGACTTC 360

Qy     361 TCAGGCTGGGCCACCCCCAGCCCCCTTGCTGCTTAAACCACTGGGGACTCGAAGTCAGAG 420
Db     361 TCAGGCTGGGCCACCCCCAGCCCCCTTGCTGCTTAAACCACTGGGGACTCGAAGTCAGAG 420

Qy     421 CTCCGTGTGATCAGGGAAGGGCTGCTTAGGAGAGGGCAGCGTCCAGGCTCTGCCAGACAT 480
Db     421 CTCCGTGTGATCAGGGAAGGGCTGCTTAGGAGAGGGCAGCGTCCAGGCTCTGCCAGACAT 480

Qy     481 CATGCTCAGGATTCTCAAGGAGGGCTGAGGGTCCCTAAGACCCCACTCCCGTGACCCAAC 540
Db     481 CATGCTCAGGATTCTCAAGGAGGGCTGAGGGTCCCTAAGACCCCACTCCCGTGACCCAAC 540

Qy     541 CCCCACTCCAATGCTCACTCCCGTGACCCAACCCCCTCTTCATTGTCATTCCAACCCCCA 600
Db     541 CCCCACTCCAATGCTCACTCCCGTGACCCAACCCCCTCTTCATTGTCATTCCAACCCCCA 600

Qy     601 CCCCACATCCCCCACCCCATCCCTCAACCCTGATGCCCATCCGCCCAGCCATTCCACCCT 660
Db     601 CCCCACATCCCCCACCCCATCCCTCAACCCTGATGCCCATCCGCCCAGCCATTCCACCCT 660

Qy     661 CACCCCCACCCCCACCCCCACGCCCACTCCCACCCCCACCCAGGCAGGATCCGGTTCCCG 720
Db     661 CACCCCCACCCCCACCCCCACGCCCACTCCCACCCCCACCCAGGCAGGATCCGGTTCCCG 720

Qy     721 CCAGGAAACATCCGGGTGCCCGGATGTGACGCCACTGACTTGCGCATTGTGGGGCAGAGA 780
Db     721 CCAGGAAACATCCGGGTGCCCGGATGTGACGCCACTGACTTGCGCATTGTGGGGCAGAGA 780

Qy     781 GAAGCGAGGTTTCCATTCTGAGGGACGGCGTAGAGTTCGGCCGAAGGAACCTGACCCAGG 840
Db     781 GAAGCGAGGTTTCCATTCTGAGGGACGGCGTAGAGTTCGGCCGAAGGAACCTGACCCAGG 840

Qy     841 CTCTGTGAGGAGGCAAGGTGAGAGGCTGAGGGAGGACTGAGGACCCCGCCACTCCAAATA 900
Db     841 CTCTGTGAGGAGGCAAGGTGAGAGGCTGAGGGAGGACTGAGGACCCCGCCACTCCAAATA 900

Qy     901 GAGAGCCCCAAATATTCCAGCCCCGCCCTTGCTGCCAGCCCTGGCCCACCCGCGGGAAGA 960
Db     901 GAGAGCCCCAAATATTCCAGCCCCGCCCTTGCTGCCAGCCCTGGCCCACCCGCGGGAAGA 960

Qy     961 CGTCTCAGCCTGGGCTGCCCCCAGACCCCTGCTCCAAAAGCCTTGAGAGACACCAGGTTC 1020
Db     961 CGTCTCAGCCTGGGCTGCCCCCAGACCCCTGCTCCAAAAGCCTTGAGAGACACCAGGTTC 1020

Qy    1021 TTCTCCCCAAGCTCTGGAATCAGAGGTTGCTGTGACCAGGGCAGGACTGGTTAGGAGAGG 1080
Db    1021 TTCTCCCCAAGCTCTGGAATCAGAGGTTGCTGTGACCAGGGCAGGACTGGTTAGGAGAGG 1080

Qy    1081 GCAGGGCACAGGCTCTGCCAGGCATCAAGATCAGCACCCAAGAGGGAGGGCTGTGGGCCC 1140
Db    1081 GCAGGGCACAGGCTCTGCCAGGCATCAAGATCAGCACCCAAGAGGGAGGGCTGTGGGCCC 1140

Qy    1141 CCAAGACTGCACTCCAATCCCCACTCCCACCCCATTCGCATTCCCATTCCCCACCCAACC 1200
Db    1141 CCAAGACTGCACTCCAATCCCCACTCCCACCCCATTCGCATTCCCATTCCCCACCCAACC 1200

Qy    1201 CCCATCTCCTCAGCTACACCTCCACCCCCATCCCTACTCCTACTCCGTCACCTGACCACC 1260
Db    1201 CCCATCTCCTCAGCTACACCTCCACCCCCATCCCTACTCCTACTCCGTCACCTGACCACC 1260

Qy    1261 ACCCTCCAGCCCCAGCACCAGCCCCAACCCTTCTGCCACCTCACCCTCACTGCCCCCAAC 1320
Db    1261 ACCCTCCAGCCCCAGCACCAGCCCCAACCCTTCTGCCACCTCACCCTCACTGCCCCCAAC 1320

Qy    1321 CCCACCCTCATCTCTCTCATGTGCCCCACTCCCATCGCCTCCCCCATTCTGGCAGAATCC 1380
Db    1321 CCCACCCTCATCTCTCTCATGTGCCCCACTCCCATCGCCTCCCCCATTCTGGCAGAATCC 1380

Qy    1381 GGTTTGCCCCTGCTCTCAACCCAGGGAAGCCCTGGTAGGCCCGATGTGAAACCACTGACT 1440
Db    1381 GGTTTGCCCCTGCTCTCAACCCAGGGAAGCCCTGGTAGGCCCGATGTGAAACCACTGACT 1440

Qy    1441 TGAACCTCACAGATCTGAGAGAAGCCAGGTTCATTTAATGGTTCTGAGGGGCGGCTTGAG 1500
Db    1441 TGAACCTCACAGATCTGAGAGAAGCCAGGTTCATTTAATGGTTCTGAGGGGCGGCTTGAG 1500

Qy    1501 ATCCACTGAGGGGAGTGGTTTTAGGCTCTGTGAGGAGGCAAGGTGAGATGCTGAGGGAGG 1560
Db    1501 ATCCACTGAGGGGAGTGGTTTTAGGCTCTGTGAGGAGGCAAGGTGAGATGCTGAGGGAGG 1560

Qy    1561 ACTGAGGAGGCACACACCCCAGGTAGATGGCCCCAAAATGATCCAGTACCACCCCTGCTG 1620
Db    1561 ACTGAGGAGGCACACACCCCAGGTAGATGGCCCCAAAATGATCCAGTACCACCCCTGCTG 1620

Qy    1621 CCAGCCCTGGACCACCCGGCCAGGACAGATGTCTCAGCTGGACCACCCCCCGTCCCGTCC 1680
Db    1621 CCAGCCCTGGACCACCCGGCCAGGACAGATGTCTCAGCTGGACCACCCCCCGTCCCGTCC 1680

Qy    1681 CACTGCCACTTAACCCACAGGGCAATCTGTAGTCATAGCTTATGTGACCGGGGCAGGGTT 1740
Db    1681 CACTGCCACTTAACCCACAGGGCAATCTGTAGTCATAGCTTATGTGACCGGGGCAGGGTT 1740

Qy    1741 GGTCAGGAGAGGCAGGGCCCAGGCATCAAGGTCCAGCATCCGCCCGGCATTAGGGTCAGG 1800
Db    1741 GGTCAGGAGAGGCAGGGCCCAGGCATCAAGGTCCAGCATCCGCCCGGCATTAGGGTCAGG 1800

Qy    1801 ACCCTGGGAGGGAACTGAGGGTTCCCCACCCACACCTGTCTCCTCATCTCCACCGCCACC 1860
Db    1801 ACCCTGGGAGGGAACTGAGGGTTCCCCACCCACACCTGTCTCCTCATCTCCACCGCCACC 1860
```

```
Qy   1861 CCACTCACATTCCCATACCTACCCCCTACCCCCAACCTCATCTTGTCAGAATCCCTGCTG 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 CCACTCACATTCCCATACCTACCCCCTACCCCCAACCTCATCTTGTCAGAATCCCTGCTG 1920

Qy   1921 TCAACCCACGGAAGCCACGGGAATGGCGGCCAGGCACTCGGATCTTGACGTCCCCATCCA 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 TCAACCCACGGAAGCCACGGGAATGGCGGCCAGGCACTCGGATCTTGACGTCCCCATCCA 1980

Qy   1981 GGGTCTGATGGAGGGAAGGGGCTTGAACAGGGCCTCAGGGGAGCAGAGGGAGGGCCCTAC 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 GGGTCTGATGGAGGGAAGGGGCTTGAACAGGGCCTCAGGGGAGCAGAGGGAGGGCCCTAC 2040

Qy   2041 TGCGAGATGAGGGAGGCCTCAGAGGACCCAGCACCCTAGGACACCGCACCCCTGTCTGAG 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 TGCGAGATGAGGGAGGCCTCAGAGGACCCAGCACCCTAGGACACCGCACCCCTGTCTGAG 2100

Qy   2101 ACTGAGGCTGCCACTTCTGGCCTCAAGAATCAGAACGATGGGGACTCAGATTGCATGGGG 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 ACTGAGGCTGCCACTTCTGGCCTCAAGAATCAGAACGATGGGGACTCAGATTGCATGGGG 2160

Qy   2161 GTGGGACCCAGGCCTGCAAGGCTTACGCGGAGGAAGAGGAGGGAGGACTCAGGGGACCTT 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 GTGGGACCCAGGCCTGCAAGGCTTACGCGGAGGAAGAGGAGGGAGGACTCAGGGGACCTT 2220

Qy   2221 GGAATCCAGATCAGTGTGGACCTCGGCCCTGAGAGGTCCAGGGCACGGTGGCCACATATG 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 GGAATCCAGATCAGTGTGGACCTCGGCCCTGAGAGGTCCAGGGCACGGTGGCCACATATG 2280

Qy   2281 GCCCATATTTCCTGCATCTTTGAGGTGACAGGACAGAGCTGTGGTCTGAGAAGTGGGGCC 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 GCCCATATTTCCTGCATCTTTGAGGTGACAGGACAGAGCTGTGGTCTGAGAAGTGGGGCC 2340

Qy   2341 TCAGGTCAACAGAGGGAGGAGTTCCAGGATCCATATGGCCCAAGATGTGCCCCCTTCATG 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 TCAGGTCAACAGAGGGAGGAGTTCCAGGATCCATATGGCCCAAGATGTGCCCCCTTCATG 2400

Qy   2401 AGGACTGGGGATATCCCCGGCTCAGAAAGAAGGGACTCCACACAGTCTGGCTGTCCCCTT 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 AGGACTGGGGATATCCCCGGCTCAGAAAGAAGGGACTCCACACAGTCTGGCTGTCCCCTT 2460

Qy   2461 TTAGTAGCTCTAGGGGGACCAGATCAGGGATGGCGGTATGTTCCATTCTCACTTGTACCA 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 TTAGTAGCTCTAGGGGGACCAGATCAGGGATGGCGGTATGTTCCATTCTCACTTGTACCA 2520

Qy   2521 CAGGCAGGAAGTTGGGGGGCCCTCAGGGAGATGGGGTCTTGGGGTAAAGGGGGGATGTCT 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 CAGGCAGGAAGTTGGGGGGCCCTCAGGGAGATGGGGTCTTGGGGTAAAGGGGGGATGTCT 2580

Qy   2581 ACTCATGTCAGGGAATTGGGGGTTGAGGAAGCACAGGCGCTGGCAGGAATAAAGATGAGT 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 ACTCATGTCAGGGAATTGGGGGTTGAGGAAGCACAGGCGCTGGCAGGAATAAAGATGAGT 2640

Qy   2641 GAGACAGACAAGGCTATTGGAATCCACACCCCAGAACCAAAGGGGTCAGCCCTGGACACC 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 GAGACAGACAAGGCTATTGGAATCCACACCCCAGAACCAAAGGGGTCAGCCCTGGACACC 2700

Qy   2701 TCACCCAGGATGTGGCTTCTTTTTCACTCCTGTTTCCAGATCTGGGGCAGGTGAGGACCT 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 TCACCCAGGATGTGGCTTCTTTTTCACTCCTGTTTCCAGATCTGGGGCAGGTGAGGACCT 2760

Qy   2761 CATTCTCAGAGGGTGACTCAGGTCAACGTAGGGACCCCCATCTGGTCTAAAGACAGAGCG 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 CATTCTCAGAGGGTGACTCAGGTCAACGTAGGGACCCCCATCTGGTCTAAAGACAGAGCG 2820

Qy   2821 GTCCCAGGATCTGCCATGCGTTCGGGTGAGGAACATGAGGGAGGACTGAGGGTACCCCAG 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 GTCCCAGGATCTGCCATGCGTTCGGGTGAGGAACATGAGGGAGGACTGAGGGTACCCCAG 2880

Qy   2881 GACCAGAACACTGAGGGAGACTGCACAGAAATCAGCCCTGCCCCTGCTGTCACCCCAGAG 2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881 GACCAGAACACTGAGGGAGACTGCACAGAAATCAGCCCTGCCCCTGCTGTCACCCCAGAG 2940

Qy   2941 AGCATGGGCTGGGCCGTCTGCCGAGGTCCTTCCGTTATCCTGGGATCATTGATGTCAGGG 3000
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941 AGCATGGGCTGGGCCGTCTGCCGAGGTCCTTCCGTTATCCTGGGATCATTGATGTCAGGG 3000

Qy   3001 ACGGGGAGGCCTTGGTCTGAGAAGGCTGCGCTCAGGTCAGTAGAGGGAGCGTCCCAGGCC 3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3001 ACGGGGAGGCCTTGGTCTGAGAAGGCTGCGCTCAGGTCAGTAGAGGGAGCGTCCCAGGCC 3060

Qy   3061 CTGCCAGGAGTCAAGGTGAGGACCAAGCGGGCACCTCACCCAGGACACATTAATTCCAAT 3120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3061 CTGCCAGGAGTCAAGGTGAGGACCAAGCGGGCACCTCACCCAGGACACATTAATTCCAAT 3120

Qy   3121 GAATTTTGATATCTCTTGCTGCCCTTCCCCAAGGACCTAGGCACGTGTGGCCAGATGTTT 3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3121 GAATTTTGATATCTCTTGCTGCCCTTCCCCAAGGACCTAGGCACGTGTGGCCAGATGTTT 3180

Qy   3181 GTCCCCTCCTGTCCTTCCATTCCTTATCATGGATGTGAACTCTTGATTTGGATTTCTCAG 3240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3181 GTCCCCTCCTGTCCTTCCATTCCTTATCATGGATGTGAACTCTTGATTTGGATTTCTCAG 3240

Qy   3241 ACCAGCAAAAGGGCAGGATCCAGGCCCTGCCAGGAAAAATATAAGGGCCCTGCGTGAGAA 3300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3241 ACCAGCAAAAGGGCAGGATCCAGGCCCTGCCAGGAAAAATATAAGGGCCCTGCGTGAGAA 3300

Qy   3301 CAGAGGGGGTCATCCACTGCATGAGAGTGGGGATGTCACAGAGTCCAGCCCACCCTCCTG 3360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3301 CAGAGGGGGTCATCCACTGCATGAGAGTGGGGATGTCACAGAGTCCAGCCCACCCTCCTG 3360

Qy   3361 GTAGCACTGAGAAGCCAGGGCTGTGCTTGCGGTCTGCACCCTGAGGGCCCGTGGATTCCT 3420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3361 GTAGCACTGAGAAGCCAGGGCTGTGCTTGCGGTCTGCACCCTGAGGGCCCGTGGATTCCT 3420

Qy   3421 CTTCCTGGAGCTCCAGGAACCAGGCAGTGAGGCCTTGGTCTGAGACAGTATCCTCAGGTC 3480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3421 CTTCCTGGAGCTCCAGGAACCAGGCAGTGAGGCCTTGGTCTGAGACAGTATCCTCAGGTC 3480

Qy   3481 ACAGAGCAGAGGATGCACAGGGTGTGCCAGCAGTGAATGTTTGCCCTGAATGCACACCAA 3540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3481 ACAGAGCAGAGGATGCACAGGGTGTGCCAGCAGTGAATGTTTGCCCTGAATGCACACCAA 3540

Qy   3541 GGGCCCCACCTGCCACAGGACACATAGGACTCCACAGAGTCTGGCCTCACCTCCCTACTG 3600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3541 GGGCCCCACCTGCCACAGGACACATAGGACTCCACAGAGTCTGGCCTCACCTCCCTACTG 3600

Qy   3601 TCAGTCCTGTAGAATCGACCTCTGCTGGCCGGCTGTACCCTGAGTACCCTCTCACTTCCT 3660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3601 TCAGTCCTGTAGAATCGACCTCTGCTGGCCGGCTGTACCCTGAGTACCCTCTCACTTCCT 3660

Qy   3661 CCTTCAGGTTTTCAGGGGACAGGCCAACCCAGAGGACAGGATTCCCTGGAGGCCACAGAG 3720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3661 CCTTCAGGTTTTCAGGGGACAGGCCAACCCAGAGGACAGGATTCCCTGGAGGCCACAGAG 3720

Qy   3721 GAGCACCAAGGAGAAGATCTGTAAGTAGGCCTTTGTTAGAGTCTCCAAGGTTCAGTTCTC 3780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3721 GAGCACCAAGGAGAAGATCTGTAAGTAGGCCTTTGTTAGAGTCTCCAAGGTTCAGTTCTC 3780

Qy   3781 AGCTGAGGCCTCTCACACACTCCCTCTCTCCCCAGGCCTGTGGGTCTTCATTGCCCAGCT 3840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3781 AGCTGAGGCCTCTCACACACTCCCTCTCTCCCCAGGCCTGTGGGTCTTCATTGCCCAGCT 3840

Qy   3841 CCTGCCCACACTCCTGCCTGCTGCCCTGACGAGAGTCATCATGTCTCTTGAGCAGAGGAG 3900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3841 CCTGCCCACACTCCTGCCTGCTGCCCTGACGAGAGTCATCATGTCTCTTGAGCAGAGGAG 3900

Qy   3901 TCTGCACTGCAAGCCTGAGGAAGCCCTTGAGGCCCAACAAGAGGCCCTGGGCCTGGTGTG 3960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3901 TCTGCACTGCAAGCCTGAGGAAGCCCTTGAGGCCCAACAAGAGGCCCTGGGCCTGGTGTG 3960

Qy   3961 TGTGCAGGCTGCCACCTCCTCCTCCTCTCCTCTGGTCCTGGGCACCCTGGAGGAGGTGCC 4020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3961 TGTGCAGGCTGCCACCTCCTCCTCCTCTCCTCTGGTCCTGGGCACCCTGGAGGAGGTGCC 4020

Qy   4021 CACTGCTGGGTCAACAGATCCTCCCCAGAGTCCTCAGGGAGCCTCCGCCTTTCCCACTAC 4080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   4021 CACTGCTGGGTCAACAGATCCTCCCCAGAGTCCTCAGGGAGCCTCCGCCTTTCCCACTAC 4080

Qy   4081 CATCAACTTCACTCGACAGAGGCAACCCAGTGAGGGTTCCAGCAGCCGTGAAGAGGAGGG 4140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4081 CATCAACTTCACTCGACAGAGGCAACCCAGTGAGGGTTCCAGCAGCCGTGAAGAGGAGGG 4140

Qy   4141 GCCAAGCACCTCTTGTATCCTGGAGTCCTTGTTCCGAGCAGTAATCACTAAGAAGGTGGC 4200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4141 GCCAAGCACCTCTTGTATCCTGGAGTCCTTGTTCCGAGCAGTAATCACTAAGAAGGTGGC 4200

Qy   4201 TGATTTGGTTGGTTTTCTGCTCCTCAAATATCGAGCCAGGGAGCCAGTCACAAAGGCAGA 4260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4201 TGATTTGGTTGGTTTTCTGCTCCTCAAATATCGAGCCAGGGAGCCAGTCACAAAGGCAGA 4260

Qy   4261 AATGCTGGAGAGTGTCATCAAAAATTACAAGCACTGTTTTCCTGAGATCTTCGGCAAAGC 4320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4261 AATGCTGGAGAGTGTCATCAAAAATTACAAGCACTGTTTTCCTGAGATCTTCGGCAAAGC 4320

Qy   4321 CTCTGAGTCCTTGCAGCTGGTCTTTGGCATTGACGTGAAGGAAGCAGACCCCACCGGCCA 4380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4321 CTCTGAGTCCTTGCAGCTGGTCTTTGGCATTGACGTGAAGGAAGCAGACCCCACCGGCCA 4380

Qy   4381 CTCCTATGTCCTTGTCACCTGCCTAGGTCTCTCCTATGATGGCCTGCTGGGTGATAATCA 4440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4381 CTCCTATGTCCTTGTCACCTGCCTAGGTCTCTCCTATGATGGCCTGCTGGGTGATAATCA 4440

Qy   4441 GATCATGCCCAAGACAGGCTTCCTGATAATTGTCCTGGTCATGATTGCAATGGAGGGCGG 4500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4441 GATCATGCCCAAGACAGGCTTCCTGATAATTGTCCTGGTCATGATTGCAATGGAGGGCGG 4500

Qy   4501 CCATGCTCCTGAGGAGGAAATCTGGGAGGAGCTGAGTGTGATGGAGGTGTATGATGGGAG 4560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4501 CCATGCTCCTGAGGAGGAAATCTGGGAGGAGCTGAGTGTGATGGAGGTGTATGATGGGAG 4560

Qy   4561 GGAGCACAGTGCCTATGGGGAGCCCAGGAAGCTGCTCACCCAAGATTTGGTGCAGGAAAA 4620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4561 GGAGCACAGTGCCTATGGGGAGCCCAGGAAGCTGCTCACCCAAGATTTGGTGCAGGAAAA 4620

Qy   4621 GTACCTGGAGTACGGCAGGTGCCGGACAGTGATCCCGCACGCTATGAGTTCCTGTGGGGT 4680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4621 GTACCTGGAGTACGGCAGGTGCCGGACAGTGATCCCGCACGCTATGAGTTCCTGTGGGGT 4680

Qy   4681 CCAAGGGCCCTCGCTGAAACCAGCTATGTGAAAGTCCTTGAGTATGTGATCAAGGTCAGT 4740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4681 CCAAGGGCCCTCGCTGAAACCAGCTATGTGAAAGTCCTTGAGTATGTGATCAAGGTCAGT 4740

Qy   4741 GCAAGAGTTCGCTTTTTCTTCCCATCCCTGCGTGAAGCAGCTTTGAGAGAGGAGGAAGAG 4800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4741 GCAAGAGTTCGCTTTTTCTTCCCATCCCTGCGTGAAGCAGCTTTGAGAGAGGAGGAAGAG 4800

Qy   4801 GGAGTCTGAGCATGAGTTGCAGCCAAGGCCAGTGGGAGGGGGACTGGGCCAGTGCACCTT 4860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4801 GGAGTCTGAGCATGAGTTGCAGCCAAGGCCAGTGGGAGGGGGACTGGGCCAGTGCACCTT 4860

Qy   4861 CCAGGGCCGCGTCCAGCAGCTTCCCCTGCCTCGTGTGACATGAGGCCCATTCTTCACTCT 4920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4861 CCAGGGCCGCGTCCAGCAGCTTCCCCTGCCTCGTGTGACATGAGGCCCATTCTTCACTCT 4920

Qy   4921 GAAGAGAGCGGTCAGTGTTCTCAGTAGTAGGTTTCTGTTCTATTGGGTGACTTGGAGATT 4980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4921 GAAGAGAGCGGTCAGTGTTCTCAGTAGTAGGTTTCTGTTCTATTGGGTGACTTGGAGATT 4980

Qy   4981 TATCTTTGTTCTCTTTTGGAATTGTTCAAATGTTTTTTTTTAAGGGATGGTTGAATGAAC 5040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4981 TATCTTTGTTCTCTTTTGGAATTGTTCAAATGTTTTTTTTTAAGGGATGGTTGAATGAAC 5040

Qy   5041 TTCAGCATCCAAGTTTATGAATGACAGCAGTCACACAGTTCTGTGTATATAGTTTAAGGG 5100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5041 TTCAGCATCCAAGTTTATGAATGACAGCAGTCACACAGTTCTGTGTATATAGTTTAAGGG 5100

Qy   5101 TAAGAGTCTTGTGTTTTATTCAGATTGGGAAATCCATTCTATTTTGTGAATTGGGATAAT 5160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5101 TAAGAGTCTTGTGTTTTATTCAGATTGGGAAATCCATTCTATTTTGTGAATTGGGATAAT 5160

Qy   5161 AACAGCAGTGGAATAAGTACTTAGAAATGTGAAAAATGAGCAGTAAAATAGATGAGATAA 5220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5161 AACAGCAGTGGAATAAGTACTTAGAAATGTGAAAAATGAGCAGTAAAATAGATGAGATAA 5220

Qy   5221 AGAACTAAAGAAATTAAGAGATAGTCAATTCTTGCCTTATACCTCAGTCTATTCTGTAAA 5280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5221 AGAACTAAAGAAATTAAGAGATAGTCAATTCTTGCCTTATACCTCAGTCTATTCTGTAAA 5280

Qy   5281 ATTTTTAAAGATATATGCATACCTGGATTTCCTTGGCTTCTTTGAGAATGTAAGAGAAAT 5340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5281 ATTTTTAAAGATATATGCATACCTGGATTTCCTTGGCTTCTTTGAGAATGTAAGAGAAAT 5340

Qy   5341 TAAATCTGAATAAAGAATTCTTCCTGTTCACTGGCTCTTTTCTTCTCCATGCACTGAGCA 5400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5341 TAAATCTGAATAAAGAATTCTTCCTGTTCACTGGCTCTTTTCTTCTCCATGCACTGAGCA 5400

Qy   5401 TCTGCTTTTTGGAAGGCCCTGGGTTAGTAGTGGAGATGCTAAGGTAAGCCAGACTCATAC 5460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5401 TCTGCTTTTTGGAAGGCCCTGGGTTAGTAGTGGAGATGCTAAGGTAAGCCAGACTCATAC 5460

Qy   5461 CCACCCATAGGGTCGTAGAGTCTAGGAGCTGCAGTCACGTAATCGAGGTGGCAAGATGTC 5520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5461 CCACCCATAGGGTCGTAGAGTCTAGGAGCTGCAGTCACGTAATCGAGGTGGCAAGATGTC 5520

Qy   5521 CTCTAAAGATGTAGGGAAAAGTGAGAGAGGGGTGAGGGTGTGGGGCTCCGGGTGAGAGTG 5580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5521 CTCTAAAGATGTAGGGAAAAGTGAGAGAGGGGTGAGGGTGTGGGGCTCCGGGTGAGAGTG 5580

Qy   5581 GTGGAGTGTCAATGCCCTGAGCTGGGGCATTTTGGGCTTTGGGAAACTGCAGTTCCTTCT 5640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5581 GTGGAGTGTCAATGCCCTGAGCTGGGGCATTTTGGGCTTTGGGAAACTGCAGTTCCTTCT 5640

Qy   5641 GGGGGAGCTGATTGTAATGATCTTGGGTGGATCC 5674
          |||||||||||||||||||||||||||||||||
Db   5641 GGGGGAGCTGATTGTAATGATCTTGGGTGGATCC 5674
```

```
RESULT   3
I36923
LOCUS       I36923       5674 bp    DNA             PAT       13-MAY-1997
DEFINITION  Sequence 8 from patent US 5612201.
ACCESSION   I36923
VERSION     I36923.1  GI:2084883
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 5674)
  AUTHORS   De Plaen,E., Boon-Falleur,T., Lethe ,B., Szikora,J., De Smet,C. and
            Chomez,P.
  TITLE     Isolated nucleic acid molecules useful in determining expression of
            a tumor rejection antigen precursor
  JOURNAL   Patent: US 5612201-A 8 18-MAR-1997;
FEATURES             Location/Qualifiers
     source          1. .5674
                     /organism="unknown"
BASE COUNT     1276 a   1644 c   1569 g   1185 t
ORIGIN
```

```
  Query Match             100.0%;  Score 5674;  DB 5;  Length 5674;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 5674;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CCCGGGGCACCACTGGCATCCCTCCCCCTACCACCCCCAATCCCTCCCTTTACGCCACCC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CCCGGGGCACCACTGGCATCCCTCCCCCTACCACCCCCAATCCCTCCCTTTACGCCACCC 60

Qy      61 ATCCAAACATCTTCACGCTCACCCCCAGCCCAAGCCAGGCAGAATCCGGTTCCACCCCTG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ATCCAAACATCTTCACGCTCACCCCCAGCCCAAGCCAGGCAGAATCCGGTTCCACCCCTG 120
```

```
Qy    121 CTCTCAACCCAGGGAAGCCCAGGTGCCCAGATGTGACGCCACTGACTTGAGCATTAGTGG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CTCTCAACCCAGGGAAGCCCAGGTGCCCAGATGTGACGCCACTGACTTGAGCATTAGTGG 180

Qy    181 TTAGAGAGAAGCGAGGTTTTCGGTCTGAGGGGCGGCTTGAGATCGGTGGAGGGAAGCGGG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TTAGAGAGAAGCGAGGTTTTCGGTCTGAGGGGCGGCTTGAGATCGGTGGAGGGAAGCGGG 240

Qy    241 CCCAGCTCTGTAAGGAGGCAAGGTGACATGCTGAGGGAGGACTGAGGACCCACTTACCCC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCCAGCTCTGTAAGGAGGCAAGGTGACATGCTGAGGGAGGACTGAGGACCCACTTACCCC 300

Qy    301 AGATAGAGGACCCCAAATAATCCCTTCATGCCAGTCCTGGACCATCTGGTGGTGGACTTC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AGATAGAGGACCCCAAATAATCCCTTCATGCCAGTCCTGGACCATCTGGTGGTGGACTTC 360

Qy    361 TCAGGCTGGGCCACCCCCAGCCCCCTTGCTGCTTAAACCACTGGGGACTCGAAGTCAGAG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TCAGGCTGGGCCACCCCCAGCCCCCTTGCTGCTTAAACCACTGGGGACTCGAAGTCAGAG 420

Qy    421 CTCCGTGTGATCAGGGAAGGGCTGCTTAGGAGAGGGCAGCGTCCAGGCTCTGCCAGACAT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CTCCGTGTGATCAGGGAAGGGCTGCTTAGGAGAGGGCAGCGTCCAGGCTCTGCCAGACAT 480

Qy    481 CATGCTCAGGATTCTCAAGGAGGGCTGAGGGTCCCTAAGACCCCACTCCCGTGACCCAAC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CATGCTCAGGATTCTCAAGGAGGGCTGAGGGTCCCTAAGACCCCACTCCCGTGACCCAAC 540

Qy    541 CCCCACTCCAATGCTCACTCCCGTGACCCAACCCCCTCTTCATTGTCATTCCAACCCCCA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CCCCACTCCAATGCTCACTCCCGTGACCCAACCCCCTCTTCATTGTCATTCCAACCCCCA 600

Qy    601 CCCCACATCCCCCACCCCATCCCTCAACCCTGATGCCCATCCGCCCAGCCATTCCACCCT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CCCCACATCCCCCACCCCATCCCTCAACCCTGATGCCCATCCGCCCAGCCATTCCACCCT 660

Qy    661 CACCCCCACCCCCACCCCCACGCCCACTCCCACCCCCACCCAGGCAGGATCCGGTTCCCG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CACCCCCACCCCCACCCCCACGCCCACTCCCACCCCCACCCAGGCAGGATCCGGTTCCCG 720

Qy    721 CCAGGAAACATCCGGGTGCCCGGATGTGACGCCACTGACTTGCGCATTGTGGGGCAGAGA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CCAGGAAACATCCGGGTGCCCGGATGTGACGCCACTGACTTGCGCATTGTGGGGCAGAGA 780

Qy    781 GAAGCGAGGTTTCCATTCTGAGGGACGGCGTAGAGTTCGGCCGAAGGAACCTGACCCAGG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GAAGCGAGGTTTCCATTCTGAGGGACGGCGTAGAGTTCGGCCGAAGGAACCTGACCCAGG 840

Qy    841 CTCTGTGAGGAGGCAAGGTGAGAGGCTGAGGGAGGACTGAGGACCCCGCCACTCCAAATA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CTCTGTGAGGAGGCAAGGTGAGAGGCTGAGGGAGGACTGAGGACCCCGCCACTCCAAATA 900

Qy    901 GAGAGCCCCAAATATTCCAGCCCCGCCCTTGCTGCCAGCCCTGGCCCACCCGCGGGAAGA 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 GAGAGCCCCAAATATTCCAGCCCCGCCCTTGCTGCCAGCCCTGGCCCACCCGCGGGAAGA 960

Qy    961 CGTCTCAGCCTGGGCTGCCCCCAGACCCCTGCTCCAAAAGCCTTGAGAGACACCAGGTTC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CGTCTCAGCCTGGGCTGCCCCCAGACCCCTGCTCCAAAAGCCTTGAGAGACACCAGGTTC 1020

Qy   1021 TTCTCCCCAAGCTCTGGAATCAGAGGTTGCTGTGACCAGGGCAGGACTGGTTAGGAGAGG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 TTCTCCCCAAGCTCTGGAATCAGAGGTTGCTGTGACCAGGGCAGGACTGGTTAGGAGAGG 1080

Qy   1081 GCAGGGCACAGGCTCTGCCAGGCATCAAGATCAGCACCCAAGAGGGAGGGCTGTGGGCCC 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 GCAGGGCACAGGCTCTGCCAGGCATCAAGATCAGCACCCAAGAGGGAGGGCTGTGGGCCC 1140

Qy   1141 CCAAGACTGCACTCCAATCCCCACTCCCACCCCATTCGCATTCCCATTCCCCACCCAACC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CCAAGACTGCACTCCAATCCCCACTCCCACCCCATTCGCATTCCCATTCCCCACCCAACC 1200

Qy   1201 CCCATCTCCTCAGCTACACCTCCACCCCCATCCCTACTCCTACTCCGTCACCTGACCACC 1260
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CCCATCTCCTCAGCTACACCTCCACCCCCATCCCTACTCCTACTCCGTCACCTGACCACC 1260

Qy   1261 ACCCTCCAGCCCCAGCACCAGCCCCAACCCTTCTGCCACCTCACCCTCACTGCCCCCAAC 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ACCCTCCAGCCCCAGCACCAGCCCCAACCCTTCTGCCACCTCACCCTCACTGCCCCCAAC 1320

Qy   1321 CCCACCCTCATCTCTCTCATGTGCCCCACTCCCATCGCCTCCCCCATTCTGGCAGAATCC 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CCCACCCTCATCTCTCTCATGTGCCCCACTCCCATCGCCTCCCCCATTCTGGCAGAATCC 1380

Qy   1381 GGTTTGCCCCCTGCTCTCAACCCAGGGAAGCCCTGGTAGGCCCGATGTGAAACCACTGACT 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 GGTTTGCCCCCTGCTCTCAACCCAGGGAAGCCCTGGTAGGCCCGATGTGAAACCACTGACT 1440

Qy   1441 TGAACCTCACAGATCTGAGAGAAGCCAGGTTCATTTAATGGTTCTGAGGGGCGGCTTGAG 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 TGAACCTCACAGATCTGAGAGAAGCCAGGTTCATTTAATGGTTCTGAGGGGCGGCTTGAG 1500

Qy   1501 ATCCACTGAGGGGAGTGGTTTTAGGCTCTGTGAGGAGGCAAGGTGAGATGCTGAGGGAGG 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 ATCCACTGAGGGGAGTGGTTTTAGGCTCTGTGAGGAGGCAAGGTGAGATGCTGAGGGAGG 1560

Qy   1561 ACTGAGGAGGCACACACCCCAGGTAGATGGCCCCAAAATGATCCAGTACCACCCCTGCTG 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 ACTGAGGAGGCACACACCCCAGGTAGATGGCCCCAAAATGATCCAGTACCACCCCTGCTG 1620

Qy   1621 CCAGCCCTGGACCACCCGGCCAGGACAGATGTCTCAGCTGGACCACCCCCCGTCCCGTCC 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 CCAGCCCTGGACCACCCGGCCAGGACAGATGTCTCAGCTGGACCACCCCCCGTCCCGTCC 1680

Qy   1681 CACTGCCACTTAACCCACAGGGCAATCTGTAGTCATAGCTTATGTGACCGGGGCAGGGTT 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 CACTGCCACTTAACCCACAGGGCAATCTGTAGTCATAGCTTATGTGACCGGGGCAGGGTT 1740

Qy   1741 GGTCAGGAGAGGCAGGGCCCAGGCATCAAGGTCCAGCATCCGCCCGGCATTAGGGTCAGG 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 GGTCAGGAGAGGCAGGGCCCAGGCATCAAGGTCCAGCATCCGCCCGGCATTAGGGTCAGG 1800

Qy   1801 ACCCTGGGAGGGAACTGAGGGTTCCCCACCCACACCTGTCTCCTCATCTCCACCGCCACC 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 ACCCTGGGAGGGAACTGAGGGTTCCCCACCCACACCTGTCTCCTCATCTCCACCGCCACC 1860

Qy   1861 CCACTCACATTCCCATACCTACCCCCTACCCCCAACCTCATCTTGTCAGAATCCCTGCTG 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 CCACTCACATTCCCATACCTACCCCCTACCCCCAACCTCATCTTGTCAGAATCCCTGCTG 1920

Qy   1921 TCAACCCACGGAAGCCACGGGAATGGCGGCCAGGCACTCGGATCTTGACGTCCCCATCCA 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 TCAACCCACGGAAGCCACGGGAATGGCGGCCAGGCACTCGGATCTTGACGTCCCCATCCA 1980

Qy   1981 GGGTCTGATGGAGGGAAGGGGCTTGAACAGGGCCTCAGGGGAGCAGAGGGAGGGCCCTAC 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 GGGTCTGATGGAGGGAAGGGGCTTGAACAGGGCCTCAGGGGAGCAGAGGGAGGGCCCTAC 2040

Qy   2041 TGCGAGATGAGGGAGGCCTCAGAGGACCCAGCACCCTAGGACACCGCACCCCTGTCTGAG 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 TGCGAGATGAGGGAGGCCTCAGAGGACCCAGCACCCTAGGACACCGCACCCCTGTCTGAG 2100

Qy   2101 ACTGAGGCTGCCACTTCTGGCCTCAAGAATCAGAACGATGGGGACTCAGATTGCATGGGG 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 ACTGAGGCTGCCACTTCTGGCCTCAAGAATCAGAACGATGGGGACTCAGATTGCATGGGG 2160

Qy   2161 GTGGGACCCAGGCCTGCAAGGCTTACGCGGAGGAAGAGGAGGGAGGACTCAGGGGACCTT 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 GTGGGACCCAGGCCTGCAAGGCTTACGCGGAGGAAGAGGAGGGAGGACTCAGGGGACCTT 2220

Qy   2221 GGAATCCAGATCAGTGTGGACCTCGGCCCTGAGAGGTCCAGGGCACGGTGGCCACATATG 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 GGAATCCAGATCAGTGTGGACCTCGGCCCTGAGAGGTCCAGGGCACGGTGGCCACATATG 2280

Qy   2281 GCCCATATTTCCTGCATCTTTGAGGTGACAGGACAGAGCTGTGGTCTGAGAAGTGGGGCC 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   2281 GCCCATATTTCCTGCATCTTTGAGGTGACAGGACAGAGCTGTGGTCTGAGAAGTGGGGCC 2340
Qy   2341 TCAGGTCAACAGAGGGAGGAGTTCCAGGATCCATATGGCCCAAGATGTGCCCCCTTCATG 2400
Db   2341 TCAGGTCAACAGAGGGAGGAGTTCCAGGATCCATATGGCCCAAGATGTGCCCCCTTCATG 2400
Qy   2401 AGGACTGGGGATATCCCCGGCTCAGAAAGAAGGGACTCCACACAGTCTGGCTGTCCCCTT 2460
Db   2401 AGGACTGGGGATATCCCCGGCTCAGAAAGAAGGGACTCCACACAGTCTGGCTGTCCCCTT 2460
Qy   2461 TTAGTAGCTCTAGGGGGACCAGATCAGGGATGGCGGTATGTTCCATTCTCACTTGTACCA 2520
Db   2461 TTAGTAGCTCTAGGGGGACCAGATCAGGGATGGCGGTATGTTCCATTCTCACTTGTACCA 2520
Qy   2521 CAGGCAGGAAGTTGGGGGGCCCTCAGGGAGATGGGGTCTTGGGGTAAAGGGGGGATGTCT 2580
Db   2521 CAGGCAGGAAGTTGGGGGGCCCTCAGGGAGATGGGGTCTTGGGGTAAAGGGGGGATGTCT 2580
Qy   2581 ACTCATGTCAGGGAATTGGGGGTTGAGGAAGCACAGGCGCTGGCAGGAATAAAGATGAGT 2640
Db   2581 ACTCATGTCAGGGAATTGGGGGTTGAGGAAGCACAGGCGCTGGCAGGAATAAAGATGAGT 2640
Qy   2641 GAGACAGACAAGGCTATTGGAATCCACACCCCAGAACCAAAGGGGTCAGCCCTGGACACC 2700
Db   2641 GAGACAGACAAGGCTATTGGAATCCACACCCCAGAACCAAAGGGGTCAGCCCTGGACACC 2700
Qy   2701 TCACCCAGGATGTGGCTTCTTTTTCACTCCTGTTTCCAGATCTGGGGCAGGTGAGGACCT 2760
Db   2701 TCACCCAGGATGTGGCTTCTTTTTCACTCCTGTTTCCAGATCTGGGGCAGGTGAGGACCT 2760
Qy   2761 CATTCTCAGAGGGTGACTCAGGTCAACGTAGGGACCCCCATCTGGTCTAAAGACAGAGCG 2820
Db   2761 CATTCTCAGAGGGTGACTCAGGTCAACGTAGGGACCCCCATCTGGTCTAAAGACAGAGCG 2820
Qy   2821 GTCCCAGGATCTGCCATGCGTTCGGGTGAGGAACATGAGGGAGGACTGAGGGTACCCCAG 2880
Db   2821 GTCCCAGGATCTGCCATGCGTTCGGGTGAGGAACATGAGGGAGGACTGAGGGTACCCCAG 2880
Qy   2881 GACCAGAACACTGAGGGAGACTGCACAGAAATCAGCCCTGCCCCTGCTGTCACCCCAGAG 2940
Db   2881 GACCAGAACACTGAGGGAGACTGCACAGAAATCAGCCCTGCCCCTGCTGTCACCCCAGAG 2940
Qy   2941 AGCATGGGCTGGGCCGTCTGCCGAGGTCCTTCCGTTATCCTGGGATCATTGATGTCAGGG 3000
Db   2941 AGCATGGGCTGGGCCGTCTGCCGAGGTCCTTCCGTTATCCTGGGATCATTGATGTCAGGG 3000
Qy   3001 ACGGGGAGGCCTTGGTCTGAGAAGGCTGCGCTCAGGTCAGTAGAGGGAGCGTCCCAGGCC 3060
Db   3001 ACGGGGAGGCCTTGGTCTGAGAAGGCTGCGCTCAGGTCAGTAGAGGGAGCGTCCCAGGCC 3060
Qy   3061 CTGCCAGGAGTCAAGGTGAGGACCAAGCGGGCACCTCACCCAGGACACATTAATTCCAAT 3120
Db   3061 CTGCCAGGAGTCAAGGTGAGGACCAAGCGGGCACCTCACCCAGGACACATTAATTCCAAT 3120
Qy   3121 GAATTTTGATATCTCTTGCTGCCCTTCCCCAAGGACCTAGGCACGTGTGGCCAGATGTTT 3180
Db   3121 GAATTTTGATATCTCTTGCTGCCCTTCCCCAAGGACCTAGGCACGTGTGGCCAGATGTTT 3180
Qy   3181 GTCCCCTCCTGTCCTTCCATTCCTTATCATGGATGTGAACTCTTGATTTGGATTTCTCAG 3240
Db   3181 GTCCCCTCCTGTCCTTCCATTCCTTATCATGGATGTGAACTCTTGATTTGGATTTCTCAG 3240
Qy   3241 ACCAGCAAAAGGGCAGGATCCAGGCCCTGCCAGGAAAAATATAAGGGCCCTGCGTGAGAA 3300
Db   3241 ACCAGCAAAAGGGCAGGATCCAGGCCCTGCCAGGAAAAATATAAGGGCCCTGCGTGAGAA 3300
Qy   3301 CAGAGGGGGTCATCCACTGCATGAGAGTGGGGATGTCACAGAGTCCAGCCCACCCTCCTG 3360
Db   3301 CAGAGGGGGTCATCCACTGCATGAGAGTGGGGATGTCACAGAGTCCAGCCCACCCTCCTG 3360
Qy   3361 GTAGCACTGAGAAGCCAGGGCTGTGCTTGCGGTCTGCACCCTGAGGGCCCGTGGATTCCT 3420
Db   3361 GTAGCACTGAGAAGCCAGGGCTGTGCTTGCGGTCTGCACCCTGAGGGCCCGTGGATTCCT 3420
```

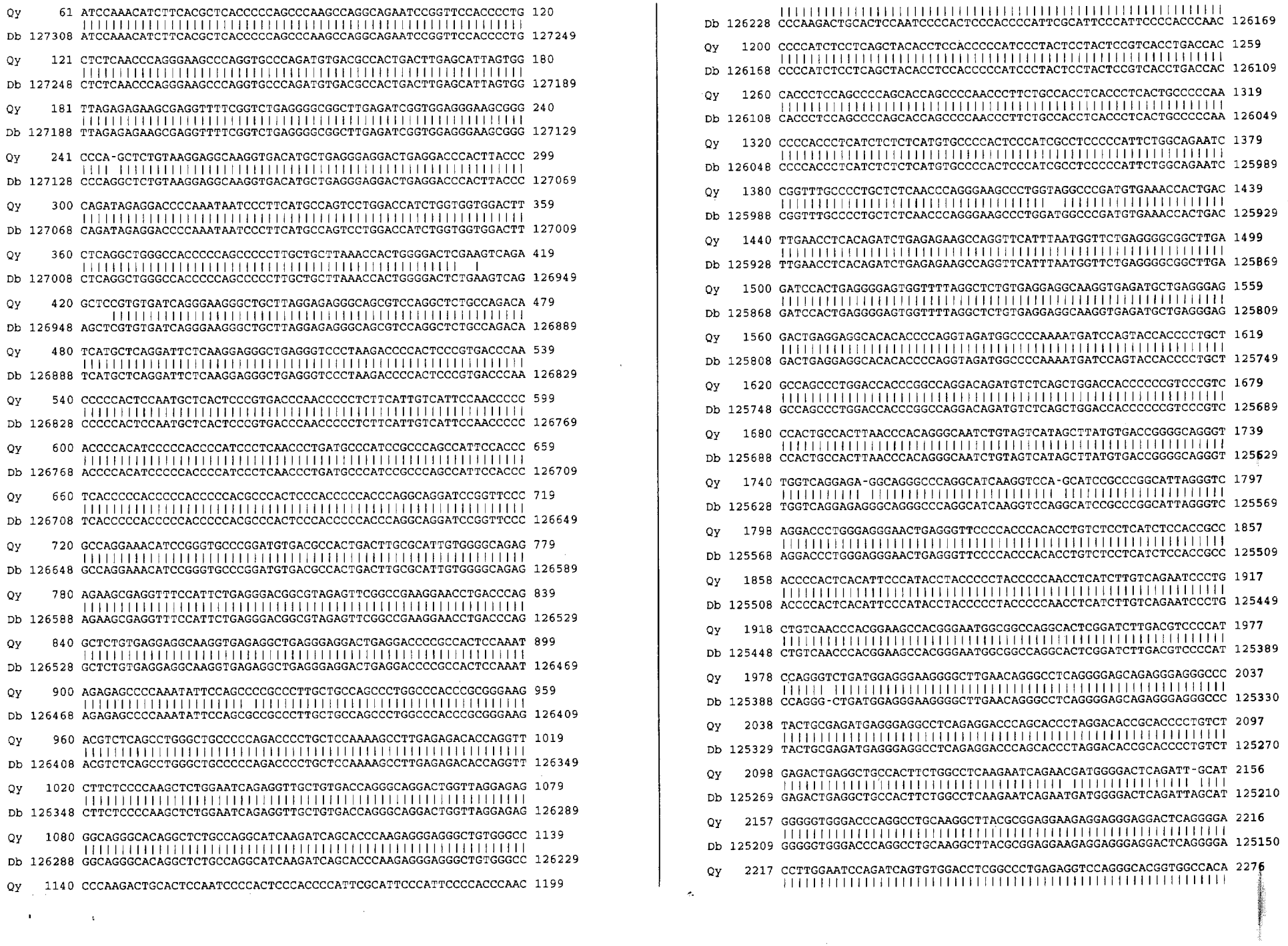
```
Qy   3421 CTTCCTGGAGCTCCAGGAACCAGGCAGTGAGGCCTTGGTCTGAGACAGTATCCTCAGGTC 3480
Db   3421 CTTCCTGGAGCTCCAGGAACCAGGCAGTGAGGCCTTGGTCTGAGACAGTATCCTCAGGTC 3480
Qy   3481 ACAGAGCAGAGGATGCACAGGGTGTGCCAGCAGTGAATGTTTGCCCTGAATGCACACCAA 3540
Db   3481 ACAGAGCAGAGGATGCACAGGGTGTGCCAGCAGTGAATGTTTGCCCTGAATGCACACCAA 3540
Qy   3541 GGGCCCCACCTGCCACAGGACACATAGGACTCCACAGAGTCTGGCCTCACCTCCCTACTG 3600
Db   3541 GGGCCCCACCTGCCACAGGACACATAGGACTCCACAGAGTCTGGCCTCACCTCCCTACTG 3600
Qy   3601 TCAGTCCTGTAGAATCGACCTCTGCTGGCCGGCTGTACCCTGAGTACCCTCTCACTTCCT 3660
Db   3601 TCAGTCCTGTAGAATCGACCTCTGCTGGCCGGCTGTACCCTGAGTACCCTCTCACTTCCT 3660
Qy   3661 CCTTCAGGTTTTCAGGGGACAGGCCAACCCAGAGGACAGGATTCCCTGGAGGCCACAGAG 3720
Db   3661 CCTTCAGGTTTTCAGGGGACAGGCCAACCCAGAGGACAGGATTCCCTGGAGGCCACAGAG 3720
Qy   3721 GAGCACCAAGGAGAAGATCTGTAAGTAGGCCTTTGTTAGAGTCTCCAAGGTTCAGTTCTC 3780
Db   3721 GAGCACCAAGGAGAAGATCTGTAAGTAGGCCTTTGTTAGAGTCTCCAAGGTTCAGTTCTC 3780
Qy   3781 AGCTGAGGCCTCTCACACACTCCCTCTCTCCCCAGGCCTGTGGGTCTTCATTGCCCAGCT 3840
Db   3781 AGCTGAGGCCTCTCACACACTCCCTCTCTCCCCAGGCCTGTGGGTCTTCATTGCCCAGCT 3840
Qy   3841 CCTGCCCACACTCCTGCCTGCTGCCCTGACGAGAGTCATCATGTCTCTTGAGCAGAGGAG 3900
Db   3841 CCTGCCCACACTCCTGCCTGCTGCCCTGACGAGAGTCATCATGTCTCTTGAGCAGAGGAG 3900
Qy   3901 TCTGCACTGCAAGCCTGAGGAAGCCCTTGAGGCCCAACAAGAGGCCCTGGGCCTGGTGTG 3960
Db   3901 TCTGCACTGCAAGCCTGAGGAAGCCCTTGAGGCCCAACAAGAGGCCCTGGGCCTGGTGTG 3960
Qy   3961 TGTGCAGGCTGCCACCTCCTCCTCCTCTCCTCTGGTCCTGGGCACCCTGGAGGAGGTGCC 4020
Db   3961 TGTGCAGGCTGCCACCTCCTCCTCCTCTCCTCTGGTCCTGGGCACCCTGGAGGAGGTGCC 4020
Qy   4021 CACTGCTGGGTCAACAGATCCTCCCCAGAGTCCTCAGGGAGCCTCCGCCTTTCCCACTAC 4080
Db   4021 CACTGCTGGGTCAACAGATCCTCCCCAGAGTCCTCAGGGAGCCTCCGCCTTTCCCACTAC 4080
Qy   4081 CATCAACTTCACTCGACAGAGGCAACCCAGTGAGGGTTCCAGCAGCCGTGAAGAGGAGGG 4140
Db   4081 CATCAACTTCACTCGACAGAGGCAACCCAGTGAGGGTTCCAGCAGCCGTGAAGAGGAGGG 4140
Qy   4141 GCCAAGCACCTCTTGTATCCTGGAGTCCTTGTTCCGAGCAGTAATCACTAAGAAGGTGGC 4200
Db   4141 GCCAAGCACCTCTTGTATCCTGGAGTCCTTGTTCCGAGCAGTAATCACTAAGAAGGTGGC 4200
Qy   4201 TGATTTGGTTGGTTTTCTGCTCCTCAAATATCGAGCCAGGGAGCCAGTCACAAAGGCAGA 4260
Db   4201 TGATTTGGTTGGTTTTCTGCTCCTCAAATATCGAGCCAGGGAGCCAGTCACAAAGGCAGA 4260
Qy   4261 AATGCTGGAGAGTGTCATCAAAAATTACAAGCACTGTTTTCCTGAGATCTTCGGCAAAGC 4320
Db   4261 AATGCTGGAGAGTGTCATCAAAAATTACAAGCACTGTTTTCCTGAGATCTTCGGCAAAGC 4320
Qy   4321 CTCTGAGTCCTTGCAGCTGGTCTTTGGCATTGACGTGAAGGAAGCAGACCCCACCGGCCA 4380
Db   4321 CTCTGAGTCCTTGCAGCTGGTCTTTGGCATTGACGTGAAGGAAGCAGACCCCACCGGCCA 4380
Qy   4381 CTCCTATGTCCTTGTCACCTGCCTAGGTCTCTCCTATGATGGCCTGCTGGGTGATAATCA 4440
Db   4381 CTCCTATGTCCTTGTCACCTGCCTAGGTCTCTCCTATGATGGCCTGCTGGGTGATAATCA 4440
Qy   4441 GATCATGCCCAAGACAGGCTTCCTGATAATTGTCCTGGTCATGATTGCAATGGAGGGCGG 4500
Db   4441 GATCATGCCCAAGACAGGCTTCCTGATAATTGTCCTGGTCATGATTGCAATGGAGGGCGG 4500
```

```
Qy   4501 CCATGCTCCTGAGGAGGAAATCTGGGAGGAGCTGAGTGTGATGGAGGTGTATGATGGGAG 4560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4501 CCATGCTCCTGAGGAGGAAATCTGGGAGGAGCTGAGTGTGATGGAGGTGTATGATGGGAG 4560

Qy   4561 GGAGCACAGTGCCTATGGGGAGCCCAGGAAGCTGCTCACCCAAGATTTGGTGCAGGAAAA 4620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4561 GGAGCACAGTGCCTATGGGGAGCCCAGGAAGCTGCTCACCCAAGATTTGGTGCAGGAAAA 4620

Qy   4621 GTACCTGGAGTACGGCAGGTGCCGGACAGTGATCCCGCACGCTATGAGTTCCTGTGGGGT 4680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4621 GTACCTGGAGTACGGCAGGTGCCGGACAGTGATCCCGCACGCTATGAGTTCCTGTGGGGT 4680

Qy   4681 CCAAGGGCCCTCGCTGAAACCAGCTATGTGAAAGTCCTTGAGTATGTGATCAAGGTCAGT 4740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4681 CCAAGGGCCCTCGCTGAAACCAGCTATGTGAAAGTCCTTGAGTATGTGATCAAGGTCAGT 4740

Qy   4741 GCAAGAGTTCGCTTTTTCTTCCCATCCCTGCGTGAAGCAGCTTTGAGAGAGGAGGAAGAG 4800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4741 GCAAGAGTTCGCTTTTTCTTCCCATCCCTGCGTGAAGCAGCTTTGAGAGAGGAGGAAGAG 4800

Qy   4801 GGAGTCTGAGCATGAGTTGCAGCCAAGGCCAGTGGGAGGGGGACTGGGCCAGTGCACCTT 4860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4801 GGAGTCTGAGCATGAGTTGCAGCCAAGGCCAGTGGGAGGGGGACTGGGCCAGTGCACCTT 4860

Qy   4861 CCAGGGCCGCGTCCAGCAGCTTCCCCTGCCTCGTGTGACATGAGGCCCATTCTTCACTCT 4920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4861 CCAGGGCCGCGTCCAGCAGCTTCCCCTGCCTCGTGTGACATGAGGCCCATTCTTCACTCT 4920

Qy   4921 GAAGAGAGCGGTCAGTGTTCTCAGTAGTAGGTTTCTGTTCTATTGGGTGACTTGGAGATT 4980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4921 GAAGAGAGCGGTCAGTGTTCTCAGTAGTAGGTTTCTGTTCTATTGGGTGACTTGGAGATT 4980

Qy   4981 TATCTTTGTTCTCTTTTGGAATTGTTCAAATGTTTTTTTTTAAGGGATGGTTGAATGAAC 5040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4981 TATCTTTGTTCTCTTTTGGAATTGTTCAAATGTTTTTTTTTAAGGGATGGTTGAATGAAC 5040

Qy   5041 TTCAGCATCCAAGTTTATGAATGACAGCAGTCACACAGTTCTGTGTATATAGTTTAAGGG 5100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5041 TTCAGCATCCAAGTTTATGAATGACAGCAGTCACACAGTTCTGTGTATATAGTTTAAGGG 5100

Qy   5101 TAAGAGTCTTGTGTTTTATTCAGATTGGGAAATCCATTCTATTTTGTGAATTGGGATAAT 5160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5101 TAAGAGTCTTGTGTTTTATTCAGATTGGGAAATCCATTCTATTTTGTGAATTGGGATAAT 5160

Qy   5161 AACAGCAGTGGAATAAGTACTTAGAAATGTGAAAAATGAGCAGTAAAATAGATGAGATAA 5220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5161 AACAGCAGTGGAATAAGTACTTAGAAATGTGAAAAATGAGCAGTAAAATAGATGAGATAA 5220

Qy   5221 AGAACTAAAGAAATTAAGAGATAGTCAATTCTTGCCTTATACCTCAGTCTATTCTGTAAA 5280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5221 AGAACTAAAGAAATTAAGAGATAGTCAATTCTTGCCTTATACCTCAGTCTATTCTGTAAA 5280

Qy   5281 ATTTTTAAAGATATATGCATACCTGGATTTCCTTGGCTTCTTTGAGAATGTAAGAGAAAT 5340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5281 ATTTTTAAAGATATATGCATACCTGGATTTCCTTGGCTTCTTTGAGAATGTAAGAGAAAT 5340

Qy   5341 TAAATCTGAATAAAGAATTCTTCCTGTTCACTGGCTCTTTTCTTCTCCATGCACTGAGCA 5400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5341 TAAATCTGAATAAAGAATTCTTCCTGTTCACTGGCTCTTTTCTTCTCCATGCACTGAGCA 5400

Qy   5401 TCTGCTTTTTGGAAGGCCCTGGGTTAGTAGTGGAGATGCTAAGGTAAGCCAGACTCATAC 5460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5401 TCTGCTTTTTGGAAGGCCCTGGGTTAGTAGTGGAGATGCTAAGGTAAGCCAGACTCATAC 5460

Qy   5461 CCACCCATAGGGTCGTAGAGTCTAGGAGCTGCAGTCACGTAATCGAGGTGGCAAGATGTC 5520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5461 CCACCCATAGGGTCGTAGAGTCTAGGAGCTGCAGTCACGTAATCGAGGTGGCAAGATGTC 5520

Qy   5521 CTCTAAAGATGTAGGGAAAAGTGAGAGAGGGGTGAGGGTGTGGGGCTCCGGGTGAGAGTG 5580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5521 CTCTAAAGATGTAGGGAAAAGTGAGAGAGGGGTGAGGGTGTGGGGCTCCGGGTGAGAGTG 5580

Qy   5581 GTGGAGTGTCAATGCCCTGAGCTGGGGCATTTTGGGCTTTGGGAAACTGCAGTTCCTTCT 5640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5581 GTGGAGTGTCAATGCCCTGAGCTGGGGCATTTTGGGCTTTGGGAAACTGCAGTTCCTTCT 5640

Qy   5641 GGGGGAGCTGATTGTAATGATCTTGGGTGGATCC 5674
          ||||||||||||||||||||||||||||||||||
Db   5641 GGGGGAGCTGATTGTAATGATCTTGGGTGGATCC 5674
```

RESULT 4
U82670/c

```
LOCUS       U82670     275159 bp    DNA             PRI       31-MAR-2000
DEFINITION  Homo sapiens chromosome Xq28 psHMG17 pseudogene, complete sequence;
            and melanoma antigen family A1 (MAGEA1) and zinc finger protein 275
            (ZNF275) genes, complete cds.
ACCESSION   U82670 U82672
VERSION     U82670.2  GI:7274891
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 275159)
  AUTHORS   Mallon,A.M., Platzer,M., Bates,R., Gloeckner,G., Botcherby,M.,
            Nordsiek,G., Strivens,M.A., Kioschis,P., Dangel,A., Cunningham,D.,
            Straw,R., Weston,P., Hunter,C., Gilbert,M., Fernando,S.,
            Goodall,K., Kerry,G., Greystrong,J.S., Clark,D., Goerdes,M.,
            Blechschmidt,K., Rump,A., Hinzmann,B., Mundy,C.R., Miller,W.,
            Poustka,A., Herman,G.E., Rhodes,M., Denny,P., Rosenthal,A. and
            Brown,S.D.M.
  TITLE     Comparative genome sequence analysis of the Bpa/Str region in mouse
            and man
  JOURNAL   Genome Res. (2000) In press
REFERENCE   2  (bases 1631 to 52906)
  AUTHORS   Gloeckner,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-DEC-1996) Genome Analysis, Institute of Molecular
            Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
REFERENCE   3  (bases 1 to 275159)
  AUTHORS   Platzer,M. and Gloeckner,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-DEC-1999) Genome Analysis, Institute of Molecular
            Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
  REMARK    Sequence update by submitter
COMMENT     On Mar 22, 2000 this sequence version replaced gi:2078526
            gi:2286115.
            This sequence is part of a larger genomic contig. The start of this
            sequence is directed towards the centromere. The start of this
            sequence (1. .250) overlaps with the end of the neighboring sequence
            deposited in GenBank Accession Number U82671.
FEATURES             Location/Qualifiers
     source          1. .275159
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /map="Xq28"
     source          <1. .35511
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="RPCIP704:L1034Q12"
     misc_feature    1. .250
                     /note="overlaps with the end of the sequence deposited in
                     GenBank Accession Number U82671"
     source          1631. .32310
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="LK0220"
     repeat_region   1900. .2216
                     /rpt_family="THE1B"
     repeat_region   complement(2590. .2993)
                     /rpt_family="MSTA"
     repeat_region   complement(2994. .3093)
                     /rpt_family="MSTA-internal"
     repeat_region   3094. .3817
```

```
                    /rpt_family="LTR8"
repeat_region       complement(3818. .5353)
                    /rpt_family="MSTA-internal"
repeat_region       complement(5354. .5746)
                    /rpt_family="MSTA"
repeat_region       5946. .6175
                    /rpt_family="HERVR"
repeat_region       6953. .7421
                    /rpt_family="HUERS-P3"
repeat_region       complement(7422. .7887)
                    /rpt_family="LTR19A"
repeat_region       7888. .7902
                    /rpt_family="HUERS-P3"
repeat_region       complement(8146. .8460)
                    /rpt_family="AluJb"
repeat_region       8603. .9022
                    /rpt_family="HUERS-P3"
repeat_region       9243. .9437
                    /rpt_family="MSTA"
repeat_region       complement(9872. .10336)
                    /rpt_family="MER41A"
repeat_region       10337. .10381
                    /rpt_family="L2"
repeat_region       complement(10383. .11111)
                    /rpt_family="MER41-internal"
repeat_region       complement(11112. .11400)
                    /rpt_family="AluSg1"
repeat_region       complement(11401. .11700)
                    /rpt_family="MER41-internal"
repeat_region       complement(11700. .13089)
                    /rpt_family="MER4-internal"
repeat_region       complement(13073. .13790)
                    /rpt_family="MER57-internal"
repeat_region       complement(13787. .13907)
                    /rpt_family="MER57-internal"
repeat_region       complement(13903. .14255)
                    /rpt_family="MER57-internal"
repeat_region       14274. .15228
                    /rpt_family="L1PA5"
repeat_region       15217. .18398
                    /rpt_family="L1PA5"
repeat_region       18442. .18975
                    /rpt_family="L1PA15"
repeat_region       19013. .19652
                    /rpt_family="LTR39"
source              19535. .52906
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="LK0813"
repeat_region       19653. .20413
                    /rpt_family="L1M3"
repeat_region       21078. .21748
                    /rpt_family="MER31-internal"
repeat_region       21752. .21886
                    /rpt_family="AluYa5/8"
repeat_region       21889. .22268
                    /rpt_family="MER31-internal"
repeat_region       22269. .22477
                    /rpt_family="AluJo"
repeat_region       22540. .22618
                    /rpt_family="MER31-internal"
repeat_region       22756. .22957
                    /rpt_family="MER31-internal"
repeat_region       22968. .23269
                    /rpt_family="AluJo"
repeat_region       23285. .23686
                    /rpt_family="L1M4"
repeat_region       complement(24864. .25145)
                    /rpt_family="AluJb"
repeat_region       25424. .25454
                    /rpt_family="(GA)n"
repeat_region       complement(25765. .25894)
                    /rpt_family="MLT1H"
misc_feature        25769. .26426
                    /note="DXS52; acc. no. GDB:187392"
repeat_region       25918. .26172
                    /rpt_family="AluY"
repeat_region       complement(26173. .26252)
                    /rpt_family="polypurine"
repeat_region       26252. .26359
                    /rpt_family="(TA)n"
repeat_region       complement(26798. .26907)
                    /rpt_family="MIR"
misc_feature        31846. .32441
                    /note="CpG island: 64.8 %GC, o/e=0.93, #CpGs=62"
repeat_region       complement(32009. .32167)
                    /rpt_family="AluJo"
repeat_region       complement(33163. .33210)
                    /rpt_family="AluJo"
repeat_region       34175. .34216
                    /rpt_family="MSTD"
repeat_region       34754. .34823
                    /rpt_family="MSTD"
repeat_region       35068. .35384
                    /rpt_family="MSTD"
repeat_region       35436. .37928
                    /rpt_family="HERVL"
repeat_region       complement(38078. .38402)
                    /rpt_family="AluY"
repeat_region       38510. .38962
                    /rpt_family="LTR16B"
repeat_region       39489. .39779
                    /rpt_family="AluSx"
repeat_region       complement(40031. .40384)
                    /rpt_family="L2"
repeat_region       complement(40510. .40802)
                    /rpt_family="L1ME1"
repeat_region       complement(40769. .40937)
                    /rpt_family="L1M4"
repeat_region       complement(41230. .41395)
                    /rpt_family="L1ME"
repeat_region       41398. .41663
                    /rpt_family="L2"
repeat_region       complement(41928. .41993)
                    /rpt_family="MLT2FB"
repeat_region       41994. .42022
                    /rpt_family="(GA)n"
repeat_region       complement(42023. .42381)
                    /rpt_family="MLT2FB"
source              42155. .103943
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="RPCI-11:K1776"
repeat_region       complement(42824. .43114)
                    /rpt_family="AluJo"
repeat_region       43685. .44167
                    /rpt_family="MLT1D"
repeat_region       44916. .45363
                    /rpt_family="MER41B"
repeat_region       47017. .47143
                    /rpt_family="MER85"
repeat_region       complement(47268. .47344)
                    /rpt_family="L2"
repeat_region       47349. .47645
                    /rpt_family="AluJb"
repeat_region       48693. .49693
                    /rpt_family="THE1C-internal"

 Query Match          98.1%;  Score 5566.2;  DB 50;  Length 275159;
 Best Local Similarity  99.6%;  Pred. No. 0;
 Matches 5654;  Conservative    0;  Mismatches   18;  Indels    7;  Gaps    7;

Qy      1 CCCGGGGCACCACTGGCATCCCTCCCCCTACCACCCCCAATCCCTCCCTTTACGCCACCC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 127368 CCCGGGGCACCACTGGCATCCCTCCCCCTACCACCCCCAATCCCTCCCTTTACGCCACCC 127309
```


Qy 61 ATCCAAACATCTTCACGCTCACCCCCAGCCCAAGCCAGGCAGAATCCGGTTCCACCCCTG 120
Db 127308 ATCCAAACATCTTCACGCTCACCCCCAGCCCAAGCCAGGCAGAATCCGGTTCCACCCCTG 127249
Qy 121 CTCTCAACCCAGGGAAGCCCAGGTGCCCAGATGTGACGCCACTGACTTGAGCATTAGTGG 180
Db 127248 CTCTCAACCCAGGGAAGCCCAGGTGCCCAGATGTGACGCCACTGACTTGAGCATTAGTGG 127189
Qy 181 TTAGAGAGAAGCGAGGTTTTCGGTCTGAGGGGCGGCTTGAGATCGGTGGAGGGAAGCGGG 240
Db 127188 TTAGAGAGAAGCGAGGTTTTCGGTCTGAGGGGCGGCTTGAGATCGGTGGAGGGAAGCGGG 127129
Qy 241 CCCA-GCTCTGTAAGGAGGCAAGGTGACATGCTGAGGGAGGACTGAGGACCCACTTACCC 299
Db 127128 CCCAGGCTCTGTAAGGAGGCAAGGTGACATGCTGAGGGAGGACTGAGGACCCACTTACCC 127069
Qy 300 CAGATAGAGGACCCCAAATAATCCCTTCATGCCAGTCCTGGACCATCTGGTGGTGGACTT 359
Db 127068 CAGATAGAGGACCCCAAATAATCCCTTCATGCCAGTCCTGGACCATCTGGTGGTGGACTT 127009
Qy 360 CTCAGGCTGGGCCACCCCCAGCCCCCTTGCTGCTTAAACCACTGGGGACTCGAAGTCAGA 419
Db 127008 CTCAGGCTGGGCCACCCCCAGCCCCCTTGCTGCTTAAACCACTGGGGACTCTGAAGTCAG 126949
Qy 420 GCTCCGTGTGATCAGGGAAGGGCTGCTTAGGAGAGGGCAGCGTCCAGGCTCTGCCAGACA 479
Db 126948 AGCTCGTGTGATCAGGGAAGGGCTGCTTAGGAGAGGGCAGCGTCCAGGCTCTGCCAGACA 126889
Qy 480 TCATGCTCAGGATTCTCAAGGAGGGCTGAGGGTCCCTAAGACCCCACTCCCGTGACCCAA 539
Db 126888 TCATGCTCAGGATTCTCAAGGAGGGCTGAGGGTCCCTAAGACCCCACTCCCGTGACCCAA 126829
Qy 540 CCCCCACTCCAATGCTCACTCCCGTGACCCAACCCCCTCTTCATTGTCATTCCAACCCCC 599
Db 126828 CCCCCACTCCAATGCTCACTCCCGTGACCCAACCCCCTCTTCATTGTCATTCCAACCCCC 126769
Qy 600 ACCCCACATCCCCCACCCCATCCCTCAACCCTGATGCCCATCCGCCCAGCCATTCCACCC 659
Db 126768 ACCCCACATCCCCCACCCCATCCCTCAACCCTGATGCCCATCCGCCCAGCCATTCCACCC 126709
Qy 660 TCACCCCCACCCCCACCCCCACGCCCACTCCCACCCCCACCCAGGCAGGATCCGGTTCCC 719
Db 126708 TCACCCCCACCCCCACCCCCACGCCCACTCCCACCCCCACCCAGGCAGGATCCGGTTCCC 126649
Qy 720 GCCAGGAAACATCCGGGTGCCCGGATGTGACGCCACTGACTTGCGCATTGTGGGGCAGAG 779
Db 126648 GCCAGGAAACATCCGGGTGCCCGGATGTGACGCCACTGACTTGCGCATTGTGGGGCAGAG 126589
Qy 780 AGAAGCGAGGTTTCCATTCTGAGGGACGGCGTAGAGTTCGGCCGAAGGAACCTGACCCAG 839
Db 126588 AGAAGCGAGGTTTCCATTCTGAGGGACGGCGTAGAGTTCGGCCGAAGGAACCTGACCCAG 126529
Qy 840 GCTCTGTGAGGAGGCAAGGTGAGAGGCTGAGGGAGGACTGAGGACCCCGCCACTCCAAAT 899
Db 126528 GCTCTGTGAGGAGGCAAGGTGAGAGGCTGAGGGAGGACTGAGGACCCCGCCACTCCAAAT 126469
Qy 900 AGAGAGCCCCAAATATTCCAGCCCCGCCCTTGCTGCCAGCCCTGGCCCACCCGCGGGAAG 959
Db 126468 AGAGAGCCCCAAATATTCCAGCGCCGCCCTTGCTGCCAGCCCTGGCCCACCCGCGGGAAG 126409
Qy 960 ACGTCTCAGCCTGGGCTGCCCCCAGACCCCTGCTCCAAAAGCCTTGAGAGACACCAGGTT 1019
Db 126408 ACGTCTCAGCCTGGGCTGCCCCCAGACCCCTGCTCCAAAAGCCTTGAGAGACACCAGGTT 126349
Qy 1020 CTTCTCCCCAAGCTCTGGAATCAGAGGTTGCTGTGACCAGGGCAGGACTGGTTAGGAGAG 1079
Db 126348 CTTCTCCCCAAGCTCTGGAATCAGAGGTTGCTGTGACCAGGGCAGGACTGGTTAGGAGAG 126289
Qy 1080 GGCAGGGCACAGGCTCTGCCAGGCATCAAGATCAGCACCCAAGAGGGAGGGCTGTGGGCC 1139
Db 126288 GGCAGGGCACAGGCTCTGCCAGGCATCAAGATCAGCACCCAAGAGGGAGGGCTGTGGGCC 126229
Qy 1140 CCCAAGACTGCACTCCAATCCCCACTCCCACCCCATTCGCATTCCCATTCCCCACCCAAC 1199
Db 126228 CCCAAGACTGCACTCCAATCCCCACTCCCACCCCATTCGCATTCCCATTCCCCACCCAAC 126169
Qy 1200 CCCCATCTCCTCAGCTACACCTCCACCCCCATCCCTACTCCTACTCCGTCACCTGACCAC 1259
Db 126168 CCCCATCTCCTCAGCTACACCTCCACCCCCATCCCTACTCCTACTCCGTCACCTGACCAC 126109
Qy 1260 CACCCTCCAGCCCCAGCACCAGCCCCAACCCTTCTGCCACCTCACCCTCACTGCCCCCAA 1319
Db 126108 CACCCTCCAGCCCCAGCACCAGCCCCAACCCTTCTGCCACCTCACCCTCACTGCCCCCAA 126049
Qy 1320 CCCCACCCTCATCTCTCTCATGTGCCCCACTCCCATCGCCTCCCCCATTCTGGCAGAATC 1379
Db 126048 CCCCACCCTCATCTCTCTCATGTGCCCCACTCCCATCGCCTCCCCCATTCTGGCAGAATC 125989
Qy 1380 CGGTTTGCCCCTGCTCTCAACCCAGGGAAGCCCTGGTAGGCCCGATGTGAAACCACTGAC 1439
Db 125988 CGGTTTGCCCCTGCTCTCAACCCAGGGAAGCCCTGGATGGCCCGATGTGAAACCACTGAC 125929
Qy 1440 TTGAACCTCACAGATCTGAGAGAAGCCAGGTTCATTTAATGGTTCTGAGGGGCGGCTTGA 1499
Db 125928 TTGAACCTCACAGATCTGAGAGAAGCCAGGTTCATTTAATGGTTCTGAGGGGCGGCTTGA 125869
Qy 1500 GATCCACTGAGGGGAGTGGTTTTAGGCTCTGTGAGGAGGCAAGGTGAGATGCTGAGGGAG 1559
Db 125868 GATCCACTGAGGGGAGTGGTTTTAGGCTCTGTGAGGAGGCAAGGTGAGATGCTGAGGGAG 125809
Qy 1560 GACTGAGGAGGCACACACCCCAGGTAGATGGCCCCAAAATGATCCAGTACCACCCCTGCT 1619
Db 125808 GACTGAGGAGGCACACACCCCAGGTAGATGGCCCCAAAATGATCCAGTACCACCCCTGCT 125749
Qy 1620 GCCAGCCCTGGACCACCCGGCCAGGACAGATGTCTCAGCTGGACCACCCCCCGTCCCGTC 1679
Db 125748 GCCAGCCCTGGACCACCCGGCCAGGACAGATGTCTCAGCTGGACCACCCCCCGTCCCGTC 125689
Qy 1680 CCACTGCCACTTAACCCACAGGGCAATCTGTAGTCATAGCTTATGTGACCGGGGCAGGGT 1739
Db 125688 CCACTGCCACTTAACCCACAGGGCAATCTGTAGTCATAGCTTATGTGACCGGGGCAGGGT 125629
Qy 1740 TGGTCAGGAGA-GGCAGGGCCCAGGCATCAAGGTCCA-GCATCCGCCCGGCATTAGGGTC 1797
Db 125628 TGGTCAGGAGAGGGCAGGGCCCAGGCATCAAGGTCCAGGCATCCGCCCGGCATTAGGGTC 125569
Qy 1798 AGGACCCTGGGAGGGAACTGAGGGTTCCCCACCCACACCTGTCTCCTCATCTCCACCGCC 1857
Db 125568 AGGACCCTGGGAGGGAACTGAGGGTTCCCCACCCACACCTGTCTCCTCATCTCCACCGCC 125509
Qy 1858 ACCCCACTCACATTCCCATACCTACCCCCTACCCCCAACCTCATCTTGTCAGAATCCCTG 1917
Db 125508 ACCCCACTCACATTCCCATACCTACCCCCTACCCCCAACCTCATCTTGTCAGAATCCCTG 125449
Qy 1918 CTGTCAACCCACGGAAGCCACGGGAATGGCGGCCAGGCACTCGGATCTTGACGTCCCCAT 1977
Db 125448 CTGTCAACCCACGGAAGCCACGGGAATGGCGGCCAGGCACTCGGATCTTGACGTCCCCAT 125389
Qy 1978 CCAGGGTCTGATGGAGGGAAGGGGCTTGAACAGGGCCTCAGGGGAGCAGAGGGAGGGCCC 2037
Db 125388 CCAGGG-CTGATGGAGGGAAGGGGCTTGAACAGGGCCTCAGGGGAGCAGAGGGAGGGCCC 125330
Qy 2038 TACTGCGAGATGAGGGAGGCCTCAGAGGACCCAGCACCCTAGGACACCGCACCCCTGTCT 2097
Db 125329 TACTGCGAGATGAGGGAGGCCTCAGAGGACCCAGCACCCTAGGACACCGCACCCCTGTCT 125270
Qy 2098 GAGACTGAGGCTGCCACTTCTGGCCTCAAGAATCAGAACGATGGGGACTCAGATT-GCAT 2156
Db 125269 GAGACTGAGGCTGCCACTTCTGGCCTCAAGAATCAGAATGATGGGGACTCAGATTAGCAT 125210
Qy 2157 GGGGGTGGGACCCAGGCCTGCAAGGCTTACGCGGAGGAAGAGGAGGGAGGACTCAGGGGA 2216
Db 125209 GGGGGTGGGACCCAGGCCTGCAAGGCTTACGCGGAGGAAGAGGAGGGAGGACTCAGGGGA 125150
Qy 2217 CCTTGGAATCCAGATCAGTGTGGACCTCGGCCCTGAGAGGTCCAGGGCACGGTGGCCACA 2276

```
Db 125149 CCTTGGAATCCAGATCAGTGTGGACCTCGGCCCTGAGAGGTCCAGGGCACGGTGGCCACA 125090

Qy   2277 TATGGCCCATATTTCCTGCATCTTTGAGGTGACAGGACAGAGCTGTGGTCTGAGAAGTGG 2336
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 125089 TATGGCCCATATTTCCTGCATCTTTGAGGTGACAGGACAGAGCTGTGGTCTGAGAAGTGG 125030

Qy   2337 GGCCTCAGGTCAACAGAGGGAGGAGTTCCAGGATCCATATGGCCCAAGATGTGCCCCCTT 2396
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 125029 GGCCTCAGGTCAACAGAGGGAGGAGTTCCAGGATCCATATGGCCCAAGATGTGCCCCCTT 124970

Qy   2397 CATGAGGACTGGGGATATCCCCGGCTCAGAAAGAAGGGACTCCACACAGTCTGGCTGTCC 2456
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 124969 CATGAGGACTGGGGATATCCCCGGCTCAGAAAGAAGGGACTCCACACAGTCTGGCTGTCC 124910

Qy   2457 CCTTTTAGTAGCTCTAGGGGGACCAGATCAGGGATGGCGGTATGTTCCATTCTCACTTGT 2516
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 124909 CCTTTTAGTAGCTCTAGGGGGACCAGATCAGGGATGGCGGTATGTTCCATTCTCACTTGT 124850

Qy   2517 ACCACAGGCAGGAAGTTGGGGGGCCCTCAGGGAGATGGGGTCTTGGGGTAAAGGGGGGAT 2576
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 124849 ACCACAGGCAGGAAGTTGGGGGGCCCTCAGGGAGATGGGGTCTTGGGGTAAAGGGGGGAT 124790

Qy   2577 GTCTACTCATGTCAGGGAATTGGGGGTTGAGGAAGCACAGGCGCTGGCAGGAATAAAGAT 2636
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 124789 GTCTACTCATGTCAGGGAATTGGGGGTTGAGGAAGCACAGGCGCTGGCAGGAATAAAGAT 124730

Qy   2637 GAGTGAGACAGACAAGGCTATTGGAATCCACACCCCAGAACCAAAGGGGTCAGCCCTGGA 2696
          ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db 124729 GAGTGAGACAGACAAGGCTATTGGAATCCTCACCCCAGAACCAAAGGGGTCAGCCCTGGA 124670

Qy   2697 CACCTCACCCAGGATGTGGCTTCTTTTTCACTCCTGTTTCCAGATCTGGGGCAGGTGAGG 2756
          ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db 124669 CACCTCACCCAGGATGTGGCTTC-TTTTCACTCCTGTTTCCAGATCTGGGGCAGGTGAGG 124611

Qy   2757 ACCTCATTCTCAGAGGGTGACTCAGGTCAACGTAGGGACCCCCATCTGGTCTAAAGACAG 2816
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 124610 ACCTCATTCTCAGAGGGTGACTCAGGTCAACGTAGGGACCCCCATCTGGTCTAAAGACAG 124551

Qy   2817 AGCGGTCCCAGGATCTGCCATGCGTTCGGGTGAGGAACATGAGGGAGGACTGAGGGTACC 2876
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 124550 AGCGGTCCCAGGATCTGCCATGCGTTCGGGTGAGGAACATGAGGGAGGACTGAGGGTACC 124491

Qy   2877 CCAGGACCAGAACACTGAGGGAGACTGCACAGAAATCAGCCCTGCCCCTGCTGTCACCCC 2936
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 124490 CCAGGACCAGAACACTGAGGGAGACTGCACAGAAATCAGCCCTGCCCCTGCTGTCACCCC 124431

Qy   2937 AGAGAGCATGGGCTGGGCCGTCTGCCGAGGTCCTTCCGTTATCCTGGGATCATTGATGTC 2996
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 124430 AGAGAGCATGGGCTGGGCCGTCTGCCGAGGTCCTTCCGTTATCCTGGGATCATTGATGTC 124371

Qy   2997 AGGGACGGGGAGGCCTTGGTCTGAGAAGGCTGCGCTCAGGTCAGTAGAGGGAGCGTCCCA 3056
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 124370 AGGGACGGGGAGGCCTTGGTCTGAGAAGGCTGCGCTCAGGTCAGTAGAGGGAGCGTCCCA 124311

Qy   3057 GGCCCTGCCAGGAGTCAAGGTGAGGACCAAGCGGGCACCTCACCCAGGACACATTAATTC 3116
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 124310 GGCCCTGCCAGGAGTCAAGGTGAGGACCAAGCGGGCACCTCACCCAGGACACATTAATTC 124251

Qy   3117 CAATGAATTTTGATATCTCTTGCTGCCCTTCCCCAAGGACCTAGGCACGTGTGGCCAGAT 3176
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 124250 CAATGAATTTTGATATCTCTTGCTGCCCTTCCCCAAGGACCTAGGCACGTGTGGCCAGAT 124191

Qy   3177 GTTTGTCCCCTCCTGTCCTTCCATTCCTTATCATGGATGTGAACTCTTGATTTGGATTTC 3236
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 124190 GTTTGTCCCCTCCTGTCCTTCCATTCCTTATCATGGATGTGAACTCTTGATTTGGATTTC 124131

Qy   3237 TCAGACCAGCAAAAGGGCAGGATCCAGGCCCTGCCAGGAAAAATATAAGGGCCCTGCGTG 3296
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 124130 TCAGACCAGCAAAAGGGCAGGATCCAGGCCCTGCCAGGAAAAATATAAGGGCCCTGCGTG 124071

Qy   3297 AGAACAGAGGGGGTCATCCACTGCATGAGAGTGGGGATGTCACAGAGTCCAGCCCACCCT 3356
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 124070 AGAACAGAGGGGGTCATCCACTGCATGAGAGTGGGGATGTCACAGAGTCCAGCCCACCCT 124011

Qy   3357 CCTGGTAGCACTGAGAAGCCAGGGCTGTGCTTGCGGTCTGCACCCTGAGGGCCCGTGGAT 3416
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 124010 CCTGGTAGCACTGAGAAGCCAGGGCTGTGCTTGCGGTCTGCACCCTGAGGGCCCGTGGAT 123951

Qy   3417 TCCTCTTCCTGGAGCTCCAGGAACCAGGCAGTGAGGCCTTGGTCTGAGACAGTATCCTCA 3476
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 123950 TCCTCTTCCTGGAGCTCCAGGAACCAGGCAGTGAGGCCTTGGTCTGAGACAGTATCCTCA 123891

Qy   3477 GGTCACAGAGCAGAGGATGCACAGGGTGTGCCAGCAGTGAATGTTTGCCCTGAATGCACA 3536
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 123890 GGTCACAGAGCAGAGGATGCACAGGGTGTGCCAGCAGTGAATGTTTGCCCTGAATGCACA 123831

Qy   3537 CCAAGGGCCCCACCTGCCACAGGACACATAGGACTCCACAGAGTCTGGCCTCACCTCCCT 3596
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 123830 CCAAGGGCCCCACCTGCCACAGGACACATAGGACTCCACAGAGTCTGGCCTCACCTCCCT 123771

Qy   3597 ACTGTCAGTCCTGTAGAATCGACCTCTGCTGGCCGGCTGTACCCTGAGTACCCTCTCACT 3656
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 123770 ACTGTCAGTCCTGTAGAATCGACCTCTGCTGGCCGGCTGTACCCTGAGTACCCTCTCACT 123711

Qy   3657 TCCTCCTTCAGGTTTTCAGGGGACAGGCCAACCCAGAGGACAGGATTCCCTGGAGGCCAC 3716
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 123710 TCCTCCTTCAGGTTTTCAGGGGACAGGCCAACCCAGAGGACAGGATTCCCTGGAGGCCAC 123651

Qy   3717 AGAGGAGCACCAAGGAGAAGATCTGTAAGTAGGCCTTTGTTAGAGTCTCCAAGGTTCAGT 3776
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 123650 AGAGGAGCACCAAGGAGAAGATCTGTAAGTAGGCCTTTGTTAGAGTCTCCAAGGTTCAGT 123591

Qy   3777 TCTCAGCTGAGGCCTCTCACACACTCCCTCTCTCCCCAGGCCTGTGGGTCTTCATTGCCC 3836
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 123590 TCTCAGCTGAGGCCTCTCACACACTCCCTCTCTCCCCAGGCCTGTGGGTCTTCATTGCCC 123531

Qy   3837 AGCTCCTGCCCACACTCCTGCCTGCTGCCCTGACGAGAGTCATCATGTCTCTTGAGCAGA 3896
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 123530 AGCTCCTGCCCACACTCCTGCCTGCTGCCCTGACGAGAGTCATCATGTCTCTTGAGCAGA 123471

Qy   3897 GGAGTCTGCACTGCAAGCCTGAGGAAGCCCTTGAGGCCCAACAAGAGGCCCTGGGCCTGG 3956
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 123470 GGAGTCTGCACTGCAAGCCTGAGGAAGCCCTTGAGGCCCAACAAGAGGCCCTGGGCCTGG 123411

Qy   3957 TGTGTGTGCAGGCTGCCACCTCCTCCTCCTCTCCTCTGGTCCTGGGCACCCTGGAGGAGG 4016
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 123410 TGTGTGTGCAGGCTGCCACCTCCTCCTCCTCTCCTCTGGTCCTGGGCACCCTGGAGGAGG 123351

Qy   4017 TGCCCACTGCTGGGTCAACAGATCCTCCCCAGAGTCCTCAGGGAGCCTCCGCCTTTCCCA 4076
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 123350 TGCCCACTGCTGGGTCAACAGATCCTCCCCAGAGTCCTCAGGGAGCCTCCGCCTTTCCCA 123291

Qy   4077 CTACCATCAACTTCACTCGACAGAGGCAACCCAGTGAGGGTTCCAGCAGCCGTGAAGAGG 4136
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 123290 CTACCATCAACTTCACTCGACAGAGGCAACCCAGTGAGGGTTCCAGCAGCCGTGAAGAGG 123231

Qy   4137 AGGGGCCAAGCACCTCTTGTATCCTGGAGTCCTTGTTCCGAGCAGTAATCACTAAGAAGG 4196
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 123230 AGGGGCCAAGCACCTCTTGTATCCTGGAGTCCTTGTTCCGAGCAGTAATCACTAAGAAGG 123171

Qy   4197 TGGCTGATTTGGTTGGTTTTCTGCTCCTCAAATATCGAGCCAGGGAGCCAGTCACAAAGG 4256
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 123170 TGGCTGATTTGGTTGGTTTTCTGCTCCTCAAATATCGAGCCAGGGAGCCAGTCACAAAGG 123111

Qy   4257 CAGAAATGCTGGAGAGTGTCATCAAAAAATTACAAGCACTGTTTTCCTGAGATCTTCGGCA 4316
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 123110 CAGAAATGCTGGAGAGTGTCATCAAAAAATTACAAGCACTGTTTTCCTGAGATCTTCGGCA 123051

Qy   4317 AAGCCTCTGAGTCCTTGCAGCTGGTCTTTGGCATTGACGTGAAGGAAGCAGACCCCACCG 4376
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 123050 AAGCCTCTGAGTCCTTGCAGCTGGTCTTTGGCATTGACGTGAAGGAAGCAGACCCCACCG 122991

Qy   4377 GCCACTCCTATGTCCTTGTCACCTGCCTAGGTCTCTCCTATGATGGCCTGCTGGGTGATA 4436
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 122990 GCCACTCCTATGTCCTTGTCACCTGCCTAGGTCTCTCCTATGATGGCCTGCTGGGTGATA 122931
```

```
Qy   4437 ATCAGATCATGCCCAAGACAGGCTTCCTGATAATTGTCCTGGTCATGATTGCAATGGAGG 4496
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 122930 ATCAGATCATGCCCAAGACAGGCTTCCTGATAATTGTCCTGGTCATGATTGCAATGGAGG 122871

Qy   4497 GCGGCCATGCTCCTGAGGAGGAAATCTGGGAGGAGCTGAGTGTGATGGAGGTGTATGATG 4556
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 122870 GCGGCCATGCTCCTGAGGAGGAAATCTGGGAGGAGCTGAGTGTGATGGAGGTGTATGATG 122811

Qy   4557 GGAGGGAGCACAGTGCCTATGGGGAGCCCAGGAAGCTGCTCACCCAAGATTTGGTGCAGG 4616
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 122810 GGAGGGAGCACAGTGCCTATGGGGAGCCCAGGAAGCTGCTCACCCAAGATTTGGTGCAGG 122751

Qy   4617 AAAAGTACCTGGAGTA-CGGCAGGTGCCGGACAGTGATCCCGCACGCTATGAGTTCCTGT 4675
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db 122750 AAAAGTACCTGGAGTACCGGCAGGTGCCGGACAGTGATCCCGCACGCTATGAGTTCCTGT 122691

Qy   4676 GGGGTCCAAGGGCCCTCGCTGAAACCAGCTATGTGAAAGTCCTTGAGTATGTGATCAAGG 4735
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 122690 GGGGTCCAAGGGCCCTCGCTGAAACCAGCTATGTGAAAGTCCTTGAGTATGTGATCAAGG 122631

Qy   4736 TCAGTGCAAGAGTTCGCTTTTTCTTCCCATCCCTGCGTGAAGCAGCTTTGAGAGAGGAGG 4795
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 122630 TCAGTGCAAGAGTTCGCTTTTTCTTCCCATCCCTGCGTGAAGCAGCTTTGAGAGAGGAGG 122571

Qy   4796 AAGAGGGAGTCTGAGCATGAGTTGCAGCCAAGGCCAGTGGGAGGGGGACTGGGCCAGTGC 4855
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 122570 AAGAGGGAGTCTGAGCATGAGTTGCAGCCAAGGCCAGTGGGAGGGGGACTGGGCCAGTGC 122511

Qy   4856 ACCTTCCAGGGCCGCGTCCAGCAGCTTCCCCTGCCTCGTGTGACATGAGGCCCATTCTTC 4915
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db 122510 ACCTTCCAGGGCCGCGTGCAGCAGCTTCCCCTGCCTCGTGTGACATGAGGCCCATTCTTC 122451

Qy   4916 ACTCTGAAGAGAGCGGTCAGTGTTCTCAGTAGTAGGTTTCTGTTCTATTGGGTGACTTGG 4975
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 122450 ACTCTGAAGAGAGCGGTCAGTGTTCTCAGTAGTAGGTTTCTGTTCTATTGGGTGACTTGG 122391

Qy   4976 AGATTTATCTTTGTTCTCTTTTGGAATTGTTCAAATGTTTTTTTTTAAGGGATGGTTGAA 5035
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 122390 AGATTTATCTTTGTTCTCTTTTGGAATTGTTCAAATGTTTTTTTTTAAGGGATGGTTGAA 122331

Qy   5036 TGAACTTCAGCATCCAAGTTTATGAATGACAGCAGTCACACAGTTCTGTGTATATAGTTT 5095
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 122330 TGAACTTCAGCATCCAAGTTTATGAATGACAGCAGTCACACAGTTCTGTGTATATAGTTT 122271

Qy   5096 AAGGGTAAGAGTCTTGTGTTTTATTCAGATTGGGAAATCCATTCTATTTTGTGAATTGGG 5155
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 122270 AAGGGTAAGAGTCTTGTGTTTTATTCAGATTGGGAAATCCATTCTATTTTGTGAATTGGG 122211

Qy   5156 ATAATAACAGCAGTGGAATAAGTACTTAGAAATGTGAAAAATGAGCAGTAAAATAGATGA 5215
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 122210 ATAATAACAGCAGTGGAATAAGTACTTAGAAATGTGAAAAATGAGCAGTAAAATAGATGA 122151

Qy   5216 GATAAAGAACTAAAGAAATTAAGAGATAGTCAATTCTTGCCTTATACCTCAGTCTATTCT 5275
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 122150 GATAAAGAACTAAAGAAATTAAGAGATAGTCAATTCTTGCCTTATACCTCAGTCTATTCT 122091

Qy   5276 GTAAAATTTTTAAAGATATATGCATACCTGGATTTCCTTGGCTTCTTTGAGAATGTAAGA 5335
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 122090 GTAAAATTTTTAAAGATATATGCATACCTGGATTTCCTTGGCTTCTTTGAGAATGTAAGA 122031

Qy   5336 GAAATTAAATCTGAATAAAGAATTCTTCCTGTTCACTGGCTCTTTTCTTCTCCATGCACT 5395
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 122030 GAAATTAAATCTGAATAAAGAATTCTTCCTGTTCACTGGCTCTTTTCTTCTCCATGCACT 121971

Qy   5396 GAGCATCTGCTTTTTGGAAGGCCCTGGGTTAGTAGTGGAGATGCTAAGGTAAGCCAGACT 5455
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121970 GAGCATCTGCTTTTTGGAAGGCCCTGGGTTAGTAGTGGAGATGCTAAGGTAAGCCAGACT 121911

Qy   5456 CATACCCACCCATAGGGTCGTAGAGTCTAGGAGCTGCAGTCACGTAATCGAGGTGGCAAG 5515
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121910 CATACCCACCCATAGGGTCGTAGAGTCTAGGAGCTGCAGTCACGTAATCGAGGTGGCAAG 121851

Qy   5516 ATGTCCTCTAAAGATGTAGGGAAAAGTGAGAGAGGGGTGAGGGTGTGGGGCTCCGGGTGA 5575
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121850 ATGTCCTCTAAAGATGTAGGGAAAAGTGAGAGAGGGGTGAGGGTGTGGGGCTCCGGGTGA 121791

Qy   5576 GAGTGGTGGAGTGTCAATGCCCTGAGCTGGGGCATTTTGGGCTTTGGGAAACTGCAGTTC 5635
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121790 GAGTGGTGGAGTGTCAATGCCCTGAGCTGGGGCATTTTGGGCTTTGGGAAACTGCAGTTC 121731

Qy   5636 CTTCTGGGGGAGCTGATTGTAATGATCTTGGGTGGATCC 5674
          |||||||||||||||||||||||||||||||||||||||
Db 121730 CTTCTGGGGGAGCTGATTGTAATGATCTTGGGTGGATCC 121692
```

```
RESULT   5
HSU82672/c
ID   HSU82672   standard; DNA; HUM; 156854 BP.
XX
AC   U82672;
XX
SV   U82672.1
XX
DT   14-MAY-1997 (Rel. 51, Created)
DT   02-JUL-1999 (Rel. 60, Last updated, Version 2)
XX
DE   Human chromosome X clone Qc15B1, complete sequence.
XX
KW   HTG.
XX
OS   Homo sapiens (human)
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Teleostomi;
OC   Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
XX
RN   [1]
RP   1-156854
RA   Gloeckner G., Rump A., Nordsiek G., Hinzmann B., Kioschis P., Heiss N.,
RA   Poustka A., Bauer D., Drescher B., Knob A., Rosenthal A.;
RT   "Sequencing and analysis of a region in Xq28 containing MAGE-1 and a
RT   putative Zinc Finger Protein";
RL   Unpublished.
XX
RN   [2]
RP   1-156854
RA   Gernot G.;
RT   ;
RL   Submitted (19-DEC-1996) to the EMBL/GenBank/DDBJ databases.
RL   Genome Analysis, IMB, Beutenbergstr. 11, Jena 07745, Germany
XX
DR   SWISS-PROT; P43355; MAG1_HUMAN.
XX
FH   Key             Location/Qualifiers
FH
FT   source          1. .156854
FT                   /chromosome="X"
FT                   /db_xref="taxon:9606"
FT                   /organism="Homo sapiens"
FT                   /map="Xq28"
FT   source          1. .40357
FT                   /db_xref="taxon:9606"
FT                   /organism="Homo sapiens"
FT                   /clone="Qc15B1"
FT   repeat_region   1057. .1295
FT                   /rpt_family="Alu-Sz"
FT   repeat_region   2504. .2666
FT                   /rpt_family="MIR"
FT   mRNA            complement(4262. .5191)
FT                   /gene="MAGE1"
FT                   /product="MAGE-1 Protein"
FT   gene            complement(4262. .5191)
FT                   /gene="MAGE1"
FT   CDS             complement(4262. .5191)
FT                   /codon_start=1
FT                   /db_xref="SWISS-PROT:P43355"
FT                   /gene="MAGE1"
FT                   /product="MAGE-1 Protein"
```

```
FT                   /protein_id="AAB54061.1"
FT                   /translation="MSLEQRSLHCKPEEALEAQQEALGLVCVQAAASSSSPLVLGTLEE
FT                   VPTAGSTDPPQSPQGASAFPTTINFTQQRQPSEGSSSREEEGPSTSCILESLFRAVITK
FT                   KVADLVGFLLLKYRAREPVTKAEMLESVIKNYKHCFPEIFGKASESLQLVFGIDVKEAD
FT                   PTGHSYVLVTCLGLSYDGLLGDNQIMPKTGFLIIVLVMIAMEGGHAPEEEIWEELSVME
FT                   VYDGREHSAYGEPRKLLTQDLVQEKYLEYRQVPDSDPARYEFLWGPRALAETSYVKVLE
FT                   YVIKVSARVRFFFPSLREAALREEEEGV"
FT   repeat_region   9623. .9892
FT                   /rpt_family="Alu-Sz"
FT   repeat_region   complement(10698. .10978)
FT                   /rpt_family="Alu-Sx"
FT   repeat_region   complement(12218. .12345)
FT                   /rpt_family="MER42C"
FT   repeat_region   complement(12636. .12851)
FT                   /rpt_family="MER42A"
FT   repeat_region   complement(12909. .13060)
FT                   /rpt_family="MER42A"
FT   repeat_region   complement(13455. .13825)
FT                   /rpt_family="MLT1A"
FT   source          13557. .53516
FT                   /db_xref="taxon:9606"
FT                   /organism="Homo sapiens"
FT                   /clone="Qc3H10"
FT   repeat_region   complement(13836. .13904)
FT                   /rpt_family="L1MC2"
FT   repeat_region   complement(13939. .14597)
FT                   /rpt_family="L1MD1"
FT   repeat_region   complement(14609. .14916)
FT                   /rpt_family="L1"
FT   repeat_region   complement(15018. .15180)
FT                   /rpt_family="L1"
FT   repeat_region   15186. .15590
FT                   /rpt_family="MLT1B"
FT   repeat_region   complement(15707. .15927)
FT                   /rpt_family="L1PA7"
FT   repeat_region   15959. .16298
FT                   /rpt_family="MLT2C2"
FT   repeat_region   complement(16377. .16753)
FT                   /rpt_family="THE1B"
FT   repeat_region   complement(16754. .18327)
FT                   /rpt_family="THE1BR"
FT   repeat_region   complement(18329. .18704)
FT                   /rpt_family="THE1B"
FT   repeat_region   complement(18983. .19802)
FT                   /rpt_family="L1PA2"
FT   repeat_region   complement(19804. .20131)
FT                   /rpt_family="L1"
FT   repeat_region   complement(20212. .20639)
FT                   /rpt_family="L1"
FT   repeat_region   complement(21048. .21637)
FT                   /rpt_family="MER11B"
FT   repeat_region   complement(21641. .21879)
FT                   /rpt_family="MER11A"
FT   repeat_region   complement(21880. .21960)
FT                   /rpt_family="MER11A"
FT   repeat_region   complement(21970. .22108)
FT                   /rpt_family="MER11A"
FT   repeat_region   24072. .24198
FT                   /rpt_family="MLT2A1"
FT   repeat_region   24341. .24389
FT                   /rpt_family="MLT2A1"
FT   repeat_region   complement(26799. .26959)
FT                   /rpt_family="MLT2B2"
FT   repeat_region   complement(27010. .27126)
FT                   /rpt_family="MLT2B2"
FT   repeat_region   27127. .27413
FT                   /rpt_family="Alu-Sg"
FT   repeat_region   complement(27433. .27667)
FT                   /rpt_family="MLT2B2"
FT   repeat_region   28489. .28775
FT                   /rpt_family="Alu-Y"
FT   repeat_region   complement(28871. .28975)
FT                   /rpt_family="MIR"
FT   repeat_region   complement(29228. .29515)
FT                   /rpt_family="Alu-Sp"
FT   repeat_region   complement(29616. .29942)
FT                   /rpt_family="L1MA10"
FT   repeat_region   complement(31359. .31633)
FT                   /rpt_family="Alu-Sq"
FT   repeat_region   complement(32548. .32808)
FT                   /rpt_family="L1"
FT   repeat_region   32809. .32857
FT                   /rpt_family="L1"
FT   repeat_region   complement(34762. .34995)
FT                   /rpt_family="MIR"
FT   source          35451. .72634
FT                   /db_xref="taxon:9606"
FT                   /organism="Homo sapiens"
FT                   /clone="Qc13B12"
FT   misc_feature    38832
FT                   /note="T deletion in Qc13B12"
FT   repeat_region   complement(41850. .41971)
FT                   /rpt_family="L1PA11"
FT   repeat_region   complement(42203. .42617)
FT                   /rpt_family="L1"
FT   repeat_region   complement(42956. .44558)
FT                   /rpt_family="TIGGER1"
FT   misc_feature    43979
FT                   /note="A deletion in Qc13B12"
FT   repeat_region   complement(44611. .45372)
FT                   /rpt_family="TIGGER1"
FT   repeat_region   46194. .46483
FT                   /rpt_family="Alu-Jo"
FT   repeat_region   46485. .46541
FT                   /rpt_family="L1MA2"
FT   repeat_region   46772. .47025
FT                   /rpt_family="MLT1F"
FT   repeat_region   48186. .48348
FT                   /rpt_family="MER5A"
FT   repeat_region   complement(48811. .48961)
FT                   /rpt_family="L1MA2"
FT   repeat_region   complement(48962. .49219)
FT                   /rpt_family="L1"
FT   repeat_region   complement(49284. .49572)
FT                   /rpt_family="Alu-Sbcg"
FT   repeat_region   complement(49578. .49626)
FT                   /rpt_family="L1"
FT   repeat_region   49645. .50522
FT                   /rpt_family="L1MA2"
FT   repeat_region   51730. .52541
FT                   /rpt_family="L1MB7"
FT   repeat_region   52733. .52811
FT                   /rpt_family="L1"
FT   repeat_region   52814. .53500
FT                   /rpt_family="L1PB3"
FT   source          63990. .107000
FT                   /db_xref="taxon:9606"
FT                   /organism="Homo sapiens"
FT                   /clone="LL0137"
FT   misc_feature    64242
FT                   /note="A in LL0137"
FT   misc_feature    66361
FT                   /note="G in LL0137"
FT   misc_feature    67333
FT                   /note="C in LL0137"
FT   misc_feature    67603
FT                   /note="G ion LL0137"
FT   misc_feature    67825
FT                   /note="A in LL0137"
FT   misc_feature    68173
FT                   /note="A in LL0137"

  Query Match             97.5%;  Score 5532.6;  DB 48;  Length 156854;
  Best Local Similarity   99.3%;  Pred. No. 0;
  Matches 5640;  Conservative    0;  Mismatches   29;  Indels   10;  Gaps    8;
```

```
Db   6911 AGGGGTGGGACCCAGGCCTGCAAGGCTTACGAGGAGGAAGAGGAGGGAGGACTCAGGGGA 6852

Qy   2217 CCTTGGAATCCAGATCAGTGTGGACCTCGGCCCTGAGAGGTCCAGGGCACGGTGGCCACA 2276
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6851 CCTTGGAATCCAGATCAGTGTGGACCTCGGCCCTGAGAGGTCCAGGGCACGGTGGCCACA 6792

Qy   2277 TATGGCCCATATTTCCTGCATCTTTGAGGTGACAGGACAGAGCTGTGGTCTGAGAAGTGG 2336
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6791 TATGGCCCATATTTCCTGCATCTTTGAGGTGACAGGACAGAGCTGTGGTCTGAGAAGTGG 6732

Qy   2337 GGCCTCAGGTCAACAGAGGGAGGAGTTCCAGGATCCATATGGCCCAAGATGTGCCCCCTT 2396
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6731 GGCCTCAGGTCAACAGAGGGAGGAGTTCCAGGATCCATATGGCCCAAGATGTGCCCCCTT 6672

Qy   2397 CATGAGGACTGGGGATATCCCCGGCTCAGAAAGAAGGGACTCCACACAGTCTGGCTGTCC 2456
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6671 CATGAGGACTGGGGATATCCCCGGCTCAGAAAGAAGGGACTCCACACAGTCTGGCTGTCC 6612

Qy   2457 CCTTTTAGTAGCTCTAGGGGGACCAGATCAGGGATGGCGGTATGTTCCATTCTCACTTGT 2516
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6611 CCTTTTAGTAGCTCTAGGGGGACCAGATCAGGGATGGCGGTATGTTCCATTCTCACTTGT 6552

Qy   2517 ACCACAGGCAGGAAGTTGGGGGGCCCTCAGGGAGATGGGGTCTTGGGGTAAAGGGGGGAT 2576
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6551 ACCACAGGCAGGAAGTTGGGGGGCCCTCAGGGAGATGGGGTCTTGGGGTAAA---GGGAT 6495

Qy   2577 GTCTACTCATGTCAGGGAATTGGGGGTTGAGGAAGCACAGGCGCTGGCAGGAATAAAGAT 2636
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6494 GTCTACTCATGTCAGGGAATTGGGGGTTGAGGAAGCACAGGCGCTGGCAGGAATAAAGAT 6435

Qy   2637 GAGTGAGACAGACAAGGCTATTGGAATCCACACCCCAGAACCAAAGGGGTCAGCCCTGGA 2696
          |||||||||||||||||||||||||||||| |||||||||||||||||| ||||||||||
Db   6434 GAGTGAGACAGACAAGGCTATTGGAATCCTCACCCCAGAACCAAAGGGGCCAGCCCTGGA 6375

Qy   2697 CACCTCACCCAGGATGTGGCTTCTTTTTCACTCCTGTTTCCAGATCTGGGGCAGGTGAGG 2756
          |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db   6374 CACCTCACCCAGGATGTGGCTTC-TTTTCACTCCTGTTTCCAGATCTGGGGCAGGTGAGG 6316

Qy   2757 ACCTCATTCTCAGAGGGTGACTCAGGTCAACGTAGGGACCCCCATCTGGTCTAAAGACAG 2816
          |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6315 ACCTCATTTTCAGAGGGTGACTCAGGTCAACGTAGGGACCCCCATCTGGTCTAAAGACAG 6256

Qy   2817 AGCGGTCCCAGGATCTGCCATGCGTTCGGGTGAGGAACATGAGGGAGGACTGAGGGTACC 2876
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db   6255 AGCGGTCCCAGGATTTGCCATGCGTTCGGGTGAGGAACATGAGGGAGGACTGAGGGTACC 6196

Qy   2877 CCAGGACCAGAACACTGAGGGAGACTGCACAGAAATCAGCCCTGCCCCTGCTGTCACCCC 2936
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6195 CCAGGACCAGAACACTGAGGGAGACTGCACAGAAATCAGCCCTGCCCCTGCTGTCACCCC 6136

Qy   2937 AGAGAGCATGGGCTGGGCCGTCTGCCGAGGTCCTTCCGTTATCCTGGGATCATTGATGTC 2996
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6135 AGAGAGCATGGGCTGGGCCGTCTGCCGAGGTCCTTCCGTTATCCTGGGATCATTGATGTC 6076

Qy   2997 AGGGACGGGGAGGCCTTGGTCTGAGAAGGCTGCGCTCAGGTCAGTAGAGGGAGCGTCCCA 3056
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6075 AGGGACGGGGAGGCCTTGGTCTGAGAAGGCTGCGCTCAGGTCAGTAGAGGGAGCGTCCCA 6016

Qy   3057 GGCCCTGCCAGGAGTCAAGGTGAGGACCAAGCGGGCACCTCACCCAGGACACATTAATTC 3116
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6015 GGCCCTGCCAGGAGTCAAGGTGAGGACCAAGCGGGCACCTCACCCAGGACACATTAATTC 5956

Qy   3117 CAATGAATTTTGATATCTCTTGCTGCCCTTCCCCAAGGACCTAGGCACGTGTGGCCAGAT 3176
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5955 CAATGAATTTTGATATCTCTTGCTGCCCTTCCCCAAGGACCTAGGCACGTGTGGCCAGAT 5896

Qy   3177 GTTTGTCCCCTCCTGTCCTTCCATTCCTTATCATGGATGTGAACTCTTGATTTGGATTTC 3236
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5895 GTTTGTCCCCTCCTGTCCTTCCATTCCTTATCATGGATGTGAACTCTTGATTTGGATTTC 5836

Qy   3237 TCAGACCAGCAAAAGGGCAGGATCCAGGCCCTGCCAGGAAAAATATAAGGGCCCTGCGTG 3296
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5835 TCAGACCAGCAAAAGGGCAGGATCCAGGCCCTGCCAGGAAAAATATAAGGGCCCTGCGTG 5776
```

```
Qy   3297 AGAACAGAGGGGGTCATCCACTGCATGAGAGTGGGGATGTCACAGAGTCCAGCCCACCCT 3356
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5775 AGAACAGAGGGGGTCATCCACTGCATGAGAGTGGGGATGTCACAGAGTCCAGCCCACCCT 5716

Qy   3357 CCTGGTAGCACTGAGAAGCCAGGGCTGTGCTTGCGGTCTGCACCCTGAGGGCCCGTGGAT 3416
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5715 CCTGGTAGCACTGAGAAGCCAGGGCTGTGCTTGCGGTCTGCACCCTGAGGGCCCGTGGAT 5656

Qy   3417 TCCTCTTCCTGGAGCTCCAGGAACCAGGCAGTGAGGCCTTGGTCTGAGACAGTATCCTCA 3476
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5655 TCCTCTTCCTGGAGCTCCAGGAACCAGGCAGTGAGGCCTTGGTCTGAGACAGTATCCTCA 5596

Qy   3477 GGTCACAGAGCAGAGGATGCACAGGGTGTGCCAGCAGTGAATGTTTGCCCTGAATGCACA 3536
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5595 GGTCACAGAGCAGAGGATGCACAGGGTGTGCCAGCAGTGAATGTTTGCCCTGAATGCACA 5536

Qy   3537 CCAAGGGCCCCCACCTGCCACAGGACACATAGGACTCCACAGAGTCTGGCCTCACCTCCCT 3596
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5535 CCAAGGGCCCCCACCTGCCACAGGACACATAGGACTCCACAGAGTCTGGCCTCACCTCCCT 5476

Qy   3597 ACTGTCAGTCCTGTAGAATCGACCTCTGCTGGCCGGCTGTACCCTGAGTACCCTCTCACT 3656
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5475 ACTGTCAGTCCTGTAGAATCGACCTCTGCTGGCCGGCTGTACCCTGAGTACCCTCTCACT 5416

Qy   3657 TCCTCCTTCAGGTTTTCAGGGGACAGGCCAACCCAGAGGACAGGATTCCCTGGAGGCCAC 3716
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5415 TCCTCCTTCAGGTTTTCAGGGGACAGGCCAACCCAGAGGACAGGATTCCCTGGAGGCCAC 5356

Qy   3717 AGAGGAGCACCAAGGAGAAGATCTGTAAGTAGGCCTTTGTTAGAGTCTCCAAGGTTCAGT 3776
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5355 AGAGGAGCACCAAGGAGAAGATCTGTAAGTAGGCCTTTGTTAGAGTCTCCAAGGTTCAGT 5296

Qy   3777 TCTCAGCTGAGGCCTCTCACACACTCCCTCTCTCCCCAGGCCTGTGGGTCTTCATTGCCC 3836
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5295 TCTCAGCTGAGGCCTCTCACACACTCCCTCTCTCCCCAGGCCTGTGGGTCTTCATTGCCC 5236

Qy   3837 AGCTCCTGCCCACACTCCTGCCTGCTGCCCCTGACGAGAGTCATCATGTCTCTTGAGCAGA 3896
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db   5235 AGCTCCTGCCCACACTCCTGCCTGTTGCCCCTGACGAGAGTCATCATGTCTCTTGAGCAGA 5176

Qy   3897 GGAGTCTGCACTGCAAGCCTGAGGAAGCCCTTGAGGCCCAACAAGAGGCCCTGGGCCTGG 3956
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5175 GGAGTCTGCACTGCAAGCCTGAGGAAGCCCTTGAGGCCCAACAAGAGGCCCTGGGCCTGG 5116

Qy   3957 TGTGTGTGCAGGCTGCCACCTCCTCCTCCTCTCCTCTGGTCCTGGGCACCCTGGAGGAGG 4016
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db   5115 TGTGTGTGCAGGCTGCCGCCTCCTCCTCCTCTCCTCTGGTCCTGGGCACCCTGGAGGAGG 5056

Qy   4017 TGCCCACTGCTGGGTCAACAGATCCTCCCCAGAGTCCTCAGGGAGCCTCCGCCTTTCCCA 4076
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5055 TGCCCACTGCTGGGTCAACAGATCCTCCCCAGAGTCCTCAGGGAGCCTCCGCCTTTCCCA 4996

Qy   4077 CTACCATCAACTTCACTCGACAGAGGCAACCCAGTGAGGGTTCCAGCAGCCGTGAAGAGG 4136
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db   4995 CTACCATCAACTTCACTCAACAGAGGCAACCCAGTGAGGGTTCCAGCAGCCGTGAAGAGG 4936

Qy   4137 AGGGGCCAAGCACCTCTTGTATCCTGGAGTCCTTGTTCCGAGCAGTAATCACTAAGAAGG 4196
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4935 AGGGGCCAAGCACCTCTTGTATCCTGGAGTCCTTGTTCCGAGCAGTAATCACTAAGAAGG 4876

Qy   4197 TGGCTGATTTGGTTGGTTTTCTGCTCCTCAAATATCGAGCCAGGGAGCCAGTCACAAAGG 4256
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4875 TGGCTGATTTGGTTGGTTTTCTGCTCCTCAAATATCGAGCCAGGGAGCCAGTCACAAAGG 4816

Qy   4257 CAGAAATGCTGGAGAGTGTCATCAAAAATTACAAGCACTGTTTTCCTGAGATCTTCGGCA 4316
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4815 CAGAAATGCTGGAGAGTGTCATCAAAAATTACAAGCACTGTTTTCCTGAGATCTTCGGCA 4756

Qy   4317 AAGCCTCTGAGTCCTTGCAGCTGGTCTTTGGCATTGACGTGAAGGAAGCAGACCCCACCG 4376
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4755 AAGCCTCTGAGTCCTTGCAGCTGGTCTTTGGCATTGACGTGAAGGAAGCAGACCCCACCG 4696
```

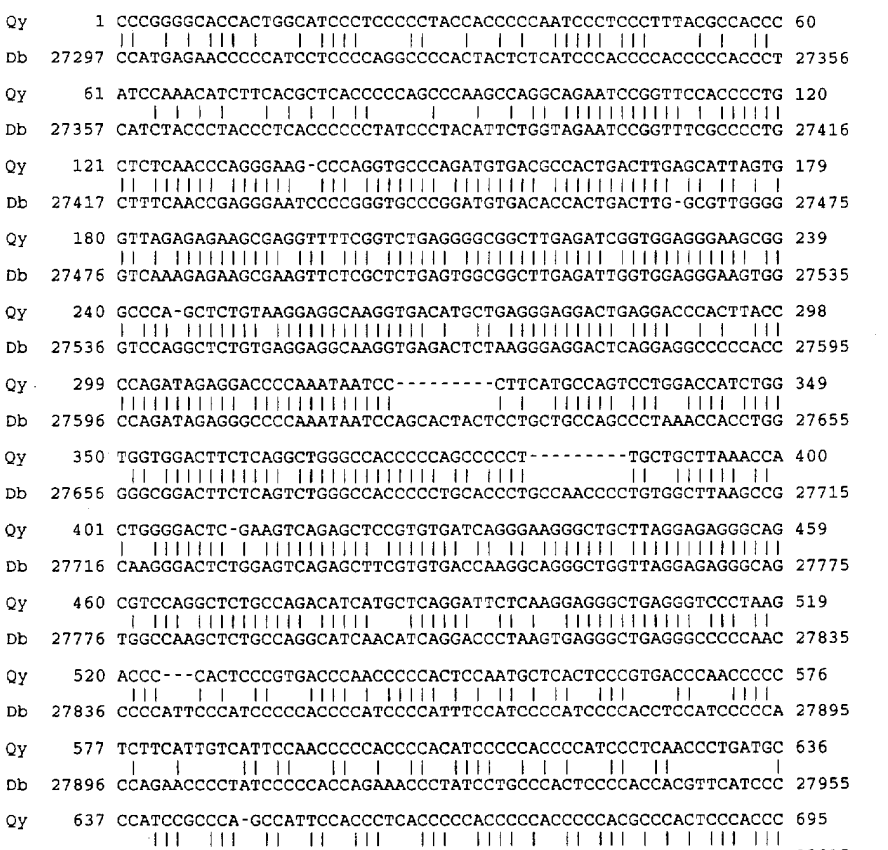
```
Qy   4377 GCCACTCCTATGTCCTTGTCACCTGCCTAGGTCTCTCCTATGATGGCCTGCTGGGTGATA 4436
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4695 GCCACTCCTATGTCCTTGTCACCTGCCTAGGTCTCTCCTATGATGGCCTGCTGGGTGATA 4636

Qy   4437 ATCAGATCATGCCCAAGACAGGCTTCCTGATAATTGTCCTGGTCATGATTGCAATGGAGG 4496
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4635 ATCAGATCATGCCCAAGACAGGCTTCCTGATAATTGTCCTGGTCATGATTGCAATGGAGG 4576

Qy   4497 GCGGCCATGCTCCTGAGGAGGAAATCTGGGAGGAGCTGAGTGTGATGGAGGTGTATGATG 4556
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4575 GCGGCCATGCTCCTGAGGAGGAAATCTGGGAGGAGCTGAGTGTGATGGAGGTGTATGATG 4516

Qy   4557 GGAGGGAGCACAGTGCCTATGGGGAGCCCAGGAAGCTGCTCACCCAAGATTTGGTGCAGG 4616
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4515 GGAGGGAGCACAGTGCCTATGGGGAGCCCAGGAAGCTGCTCACCCAAGATTTGGTGCAGG 4456

Qy   4617 AAAAGTACCTGGAGTA-CGGCAGGTGCCGGACAGTGATCCCGCACGCTATGAGTTCCTGT 4675
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   4455 AAAAGTACCTGGAGTACCGGCAGGTGCCGGACAGTGATCCCGCACGCTATGAGTTCCTGT 4396

Qy   4676 GGGGTCCAAGGGCCCTCGCTGAAACCAGCTATGTGAAAGTCCTTGAGTATGTGATCAAGG 4735
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4395 GGGGTCCAAGGGCCCTCGCTGAAACCAGCTATGTGAAAGTCCTTGAGTATGTGATCAAGG 4336

Qy   4736 TCAGTGCAAGAGTTCGCTTTTTCTTCCCATCCCTGCGTGAAGCAGCTTTGAGAGAGGAGG 4795
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4335 TCAGTGCAAGAGTTCGCTTTTTCTTCCCATCCCTGCGTGAAGCAGCTTTGAGAGAGGAGG 4276

Qy   4796 AAGAGGGAGTCTGAGCATGAGTTGCAGCCAAGGCCAGTGGGAGGGGGACTGGGCCAGTGC 4855
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4275 AAGAGGGAGTCTGAGCATGAGTTGCAGCCAAGGCCAGTGGGAGGGGGACTGGGCCAGTGC 4216

Qy   4856 ACCTTCCAGGGCCGCGTCCAGCAGCTTCCCCTGCCTCGTGTGACATGAGGCCCATTCTTC 4915
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db   4215 ACCTTCCAGGGCCGCGTGCAGCAGCTTCCCCTGCCTCGTGTGACATGAGGCCCATTCTTC 4156

Qy   4916 ACTCTGAAGAGAGCGGTCAGTGTTCTCAGTAGTAGGTTTCTGTTCTATTGGGTGACTTGG 4975
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   4155 ACTCTGAAGAGAGCGGTCAGTGTTCTCAGTAGTAGGTTTCTGTTCTGTTGGGTGACTTGG 4096

Qy   4976 AGATTTATCTTTGTTCTCTTTTGGAATTGTTCAAATGTTTTTTTTTAAGGGATGGTTGAA 5035
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4095 AGATTTATCTTTGTTCTCTTTTGGAATTGTTCAAATGTTTTTTTTTAAGGGATGGTTGAA 4036

Qy   5036 TGAACTTCAGCATCCAAGTTTATGAATGACAGCAGTCACACAGTTCTGTGTATATAGTTT 5095
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4035 TGAACTTCAGCATCCAAGTTTATGAATGACAGCAGTCACACAGTTCTGTGTATATAGTTT 3976

Qy   5096 AAGGGTAAGAGTCTTGTGTTTTATTCAGATTGGGAAATCCATTCTATTTTGTGAATTGGG 5155
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3975 AAGGGTAAGAGTCTTGTGTTTTATTCAGATTGGGAAATCCATTCTATTTTGTGAATTGGG 3916

Qy   5156 ATAATAACAGCAGTGGAATAAGTACTTAGAAATGTGAAAAATGAGCAGTAAAATAGATGA 5215
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3915 ATAATAACAGCAGTGGAATAAGTACTTAGAAATGTGAAAAATGAGCAGTAAAATAGATGA 3856

Qy   5216 GATAAAGAACTAAAGAAATTAAGAGATAGTCAATTCTTGCCTTATACCTCAGTCTATTCT 5275
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3855 GATAAAGAACTAAAGAAATTAAGAGATAGTCAATTCTTGCCTTATACCTCAGTCTATTCT 3796

Qy   5276 GTAAAATTTTTAAAGATATATGCATACCTGGATTTCCTTGGCTTCTTTGAGAATGTAAGA 5335
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3795 GTAAAATTTTTAAAGATATATGCATACCTGGATTTCCTTGGCTTCTTTGAGAATGTAAGA 3736

Qy   5336 GAAATTAAATCTGAATAAAGAATTCTTCCTGTTCACTGGCTCTTTTCTTCTCCATGCACT 5395
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3735 GAAATTAAATCTGAATAAAGAATTCTTCCTGTTCACTGGCTCTTTTCTTCTCCATGCACT 3676

Qy   5396 GAGCATCTGCTTTTTGGAAGGCCCTGGGTTAGTAGTGGAGATGCTAAGGTAAGCCAGACT 5455
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3675 GAGCATCTGCTTTTTGGAAGGCCCTGGGTTAGTAGTGGAGATGCTAAGGTAAGCCAGACT 3616

Qy   5456 CATACCCACCCATAGGGTCGTAGAGTCTAGGAGCTGCAGTCACGTAATCGAGGTGGCAAG 5515
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3615 CATACCCACCCATAGGGTCGTAGAGTCTAGGAGCTGCAGTCACGTAATCGAGGTGGCAAG 3556

Qy   5516 ATGTCCTCTAAAGATGTAGGGAAAAGTGAGAGAGGGGTGAGGGTGTGGGGCTCCGGGTGA 5575
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3555 ATGTCCTCTAAAGATGTAGGGAAAAGTGAGAGAGGGGTGAGGGTGTGGGGCTCCGGGTGA 3496

Qy   5576 GAGTGGTGGAGTGTCAATGCCCTGAGCTGGGGCATTTTGGGCTTTGGGAAACTGCAGTTC 5635
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3495 GAGTGGTGGAGTGTCAATGCCCTGAGCTGGGGCATTTTGGGCTTTGGGAAACTGCAGTTC 3436

Qy   5636 CTTCTGGGGGAGCTGATTGTAATGATCTTGGGTGGATCC 5674
          |||||||||||||||||||||||||||||||||||||   
Db   3435 CTTCTGGGGGAGCTGATTGTAATGATCTTGGGTGGATCC 3397
```

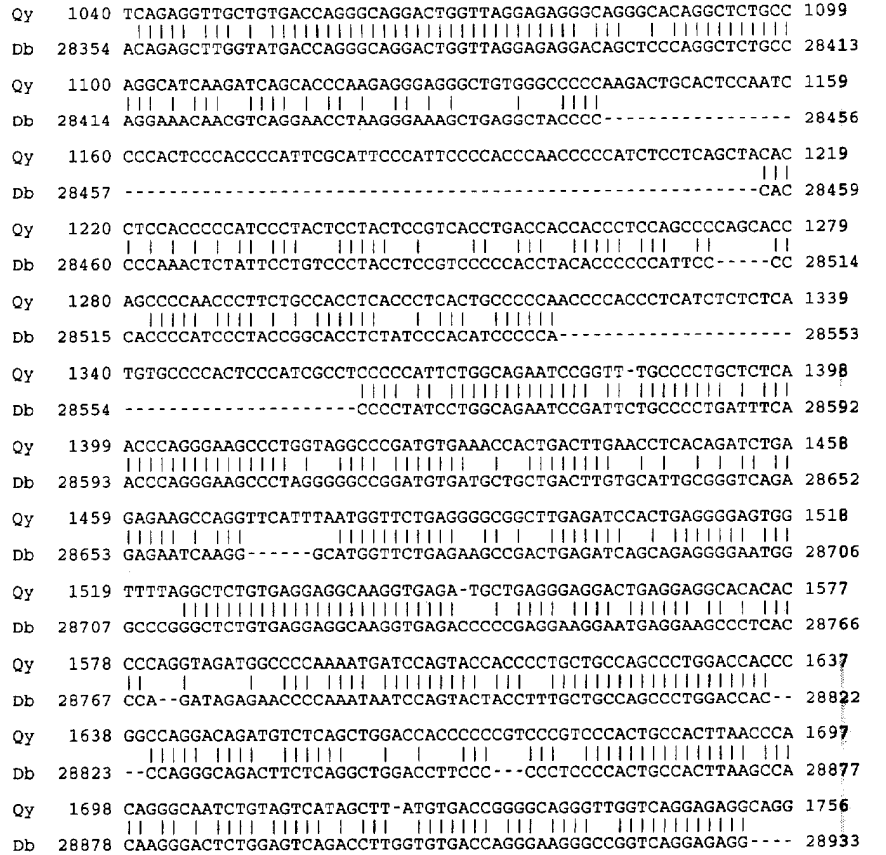
RESULT 6
AF134576
LOCUS AF134576 49375 bp DNA PRI 06-APR-1999
DEFINITION Homo sapiens MAGE10 and undefined MAGE genes, complete sequence.
ACCESSION AF134576
VERSION AF134576.1 GI:4567140
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 49375)
AUTHORS Gloeckner,G. and Rosenthal,A.
TITLE Sequence of three cosmids of human Xq28 (16c3, 3g5, 3h5) containing MAGE genes
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 49375)
AUTHORS Gloeckner,G., Rosenthal,A. and Schattevoy,R.
TITLE Direct Submission
JOURNAL Submitted (12-MAR-1999) Genome Analysis, Institute for Molecular Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany

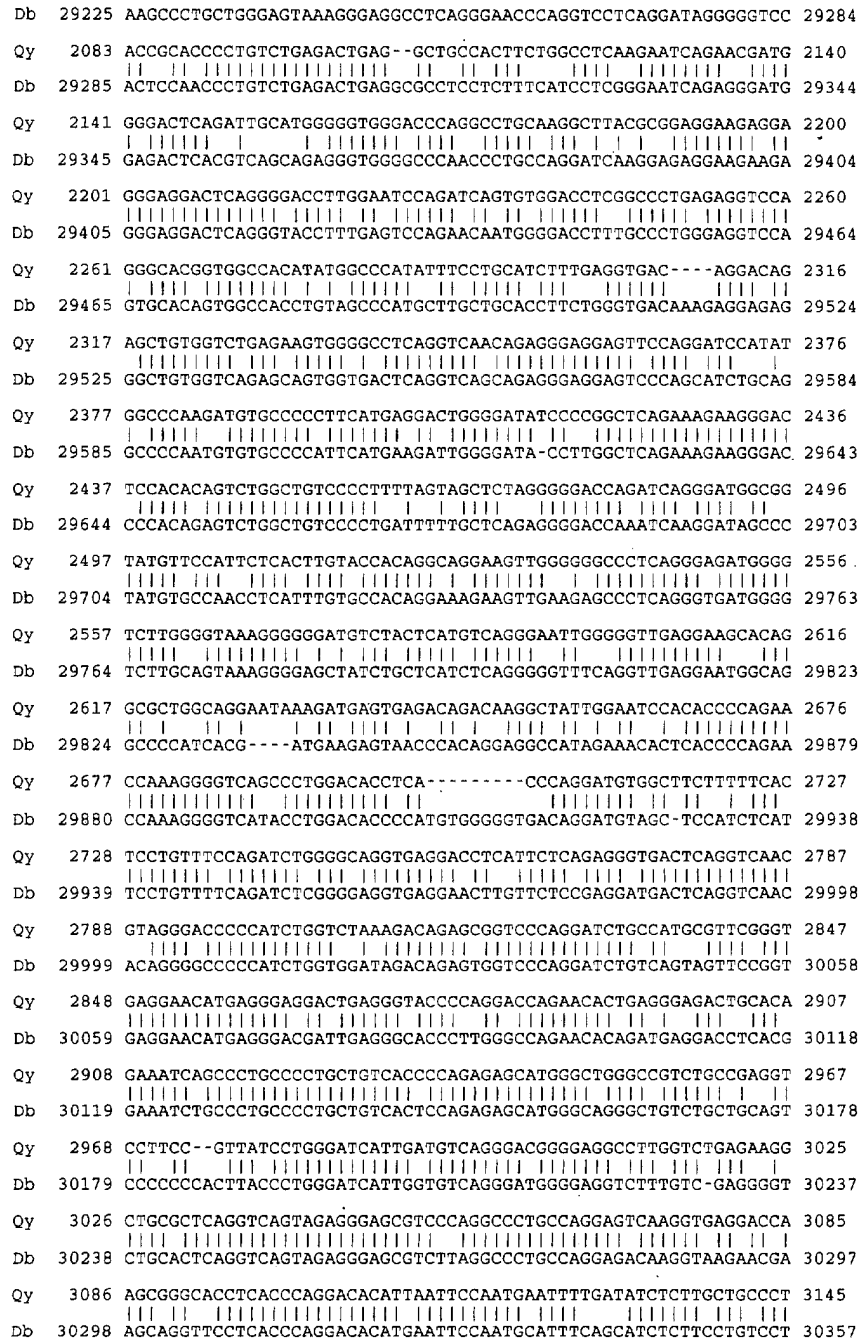
FEATURES	Location/Qualifiers
source	1. .49375
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/map="Xq28"
	/chromosome="X"
	/clone="cosmids 16c3; 3g5; 3h5"
misc_feature	1. .10000
	/note="incomplete at 3' end"
	/note="Region: Cosmid 16c3"
repeat_region	complement(10. .202)
	/rpt_family="MER93"
	/evidence=not_experimental
repeat_region	289. .372
	/rpt_family="L1M1"
	/evidence=not_experimental
repeat_region	446. .726
	/rpt_family="L1M1"
	/evidence=not_experimental
misc_feature	528. .43599
	/note="Region: Cosmid 3g5"
repeat_region	731. .4392
	/rpt_family="L1MC3"
	/evidence=not_experimental
repeat_region	complement(4393. .4696)
	/rpt_family="MER2"
	/evidence=not_experimental
repeat_region	complement(4697. .4755)
	/rpt_family="L1MA2"
	/evidence=not_experimental
repeat_region	complement(4756. .4803)
	/rpt_family="MER2"
	/evidence=not_experimental
repeat_region	4805. .4847
	/rpt_family="L1MC3"
	/evidence=not_experimental

```
repeat_region   4837. .5414
                /rpt_family="L1MC5"
                /evidence=not_experimental
repeat_region   5566. .6545
                /rpt_family="L1M4"
                /evidence=not_experimental
repeat_region   complement(6905. .6963)
                /rpt_family="MIR"
                /evidence=not_experimental
repeat_region   7671. .7697
                /note="1I1 with 85% homology to 1I2"
                /rpt_type=inverted
                /evidence=not_experimental
repeat_region   7814. .7981
                /note="homology = 66.10%, score = 48, counts = 28"
                /rpt_type=tandem
                /rpt_unit=ccccac
                /evidence=not_experimental
misc_feature    7814. .8708
                /note="GC score = 44.00 (895bp)"
                /note="Region: GC content"
                /evidence=not_experimental
misc_feature    8294. .8511
                /note="CpG island score = 0.74, GC = 66.50%, CpGs = 17"
                /note="Region: CpG island"
                /evidence=not_experimental
repeat_region   8854. .8880
                /note="1I2 with 85% homology to 1I1"
                /rpt_type=inverted
                /evidence=not_experimental
gene            9050. .11121
                /gene="MAGE10"
                /note="1. .2072 of gb|U10685|HSU10685 Human
                MAGE-10 antigen (MAGE10) gene, complete cds. 6/95 P = 0.0
                S = 9975"
misc_feature    9936. .10064
                /note="GC score = 5.60 (129bp)"
                /note="Region: GC content"
                /evidence=not_experimental
misc_feature    10412. .10533
                /note="GC score = 8.80 (122bp)"
                /note="Region: GC content"
                /evidence=not_experimental
misc_feature    11591. .49375
                /note="Region: Cosmid 3h5"
repeat_region   12154. .12482
                /rpt_family="L1MA5"
                /evidence=not_experimental
repeat_region   complement(12670. .12980)
                /rpt_family="AluSx"
                /evidence=not_experimental
repeat_region   complement(14281. .14525)
                /rpt_family="MIR"
                /evidence=not_experimental
repeat_region   complement(16047. .17011)
                /rpt_family="LTR5"
                /evidence=not_experimental
repeat_region   17625. .17747
                /rpt_family="L1ME1"
                /evidence=not_experimental
repeat_region   17800. .17815
                /rpt_family="L1M4"
                /evidence=not_experimental
repeat_region   17843. .18317
                /rpt_family="L1M4"
                /evidence=not_experimental
repeat_region   20010. .20829
                /rpt_family="L1MC/D"
                /evidence=not_experimental
repeat_region   complement(20859. .21195)
                /rpt_family="MLT2E"
                /evidence=not_experimental
repeat_region   21198. .21305
                /rpt_family="L1MC/D"
                /evidence=not_experimental
repeat_region   complement(21306. .21496)
                /rpt_family="MLT1B"
                /evidence=not_experimental
repeat_region   complement(21670. .21839)
                /rpt_family="MLT1B"
                /evidence=not_experimental
repeat_region   21844. .21992
                /rpt_family="L1MC/D"
                /evidence=not_experimental
repeat_region   22019. .22814
                /rpt_family="L1MC/D"
                /evidence=not_experimental
repeat_region   22796. .23134
                /rpt_family="L1M4"
                /evidence=not_experimental
repeat_region   23127. .23300
                /rpt_family="L1MD3"
                /evidence=not_experimental
repeat_region   23271. .23837
                /rpt_family="L1MC5"
                /evidence=not_experimental
repeat_region   24267. .25020
                /rpt_family="L1M4"
                /evidence=not_experimental
misc_feature    26071. .28617
                /note="GC score = 31.80 (2547bp)"
                /note="Region: GC content"
                /evidence=not_experimental
repeat_region   26655. .26726
                /note="homology = 79.20%, score = 36, counts = 12"
                /rpt_type=tandem
                /rpt_unit=catccc
                /evidence=not_experimental
repeat_region   27233. .27358
                /note="homology = 65.90%, score = 34, counts = 21"
                /rpt_type=tandem
                /rpt_unit=ccccca
                /evidence=not_experimental
misc_feature    27403. .27510
                /note="CpG island score = 1.29, GC = 61.10%, CpGs = 9"
                /note="Region: CpG island"
                /evidence=not_experimental
repeat_region   27829. .27906
                /note="homology = 79.50%, score = 40, counts = 13"
                /rpt_type=tandem
                /rpt_unit=tcccca
                /evidence=not_experimental
repeat_region   28990. .29073
                /note="homology = 69.00%, score = 26, counts = 14"
                /rpt_type=tandem
                /rpt_unit=ccccaa
                /evidence=not_experimental
misc_feature    30461. .30620
                /note="GC score = 5.00 (160bp)"
                /note="Region: GC content"
                /evidence=not_experimental
repeat_region   30896. .30958
                /note="homology = 82.50%, score = 20, counts = 3"
                /rpt_type=tandem
                /rpt_unit=aggctctagaggatcaccagg
                /evidence=not_experimental
gene            31110. .31778
                /gene="undefined MAGE"
                /note="1. .185 of gbp|D32077|HUMMAGEC_1 Homo
                sapiens Human mRNA for MAGE-41 protein, complete P =
                6.0e-95 S = 592-193. .283 of gbp|D32076|HUMMAGEB_1 Homo
                sapiens Human mRNA for MAGE-6 potein, complete cds; P =
                1.5e-80 S = 356"
misc_feature    31172. .31331
                /note="GC score = 13.00 (160bp)"
```

Query Match 48.9%; Score 2777.2; DB 39; Length 49375;
Best Local Similarity 74.9%; Pred. No. 0;
Matches 4355; Conservative 0; Mismatches 1109; Indels 353; Gaps 56;



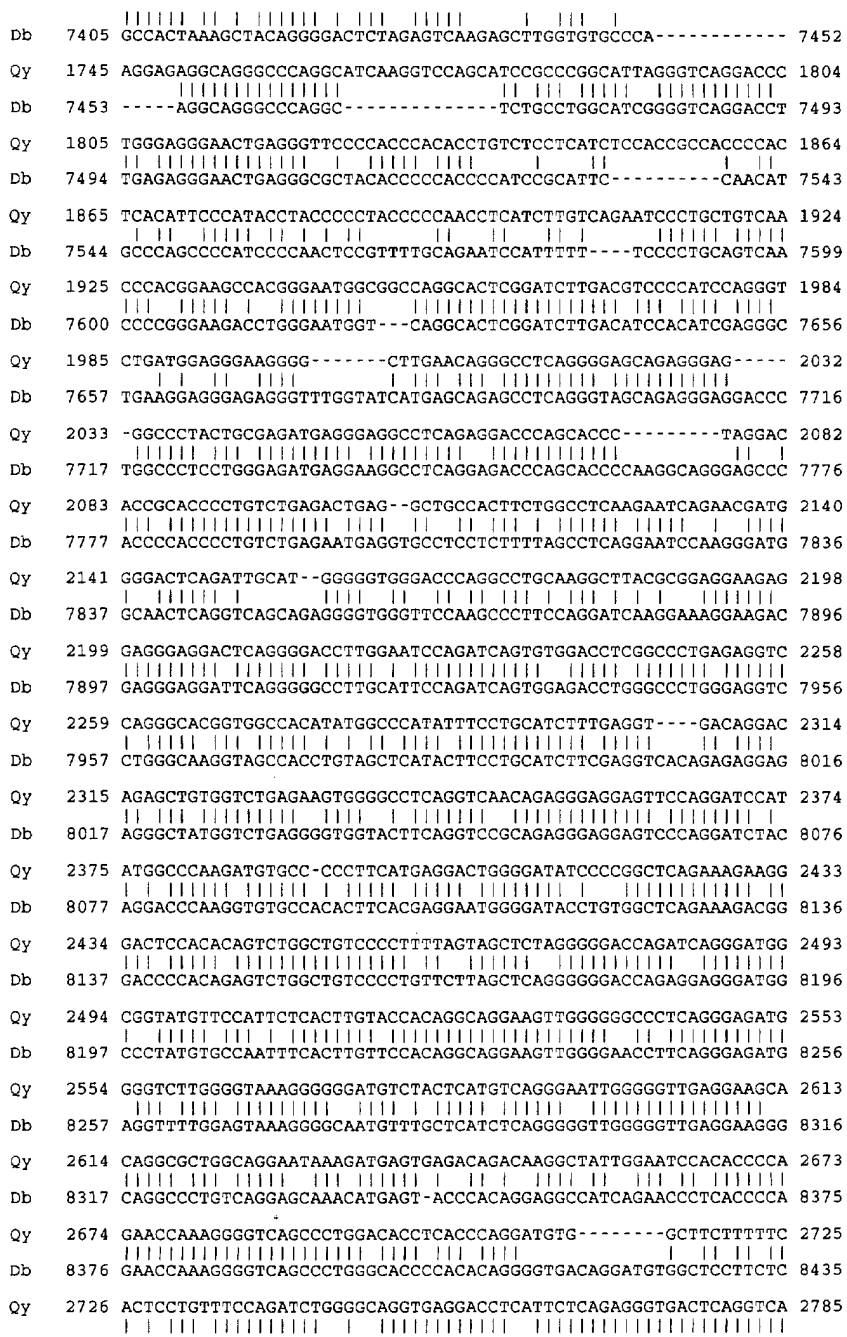



Db 29225 AAGCCCTGCTGGGAGTAAAGGGAGGCCTCAGGGAACCCAGGTCCTCAGGATAGGGGGTCC 29284
Qy 2083 ACCGCACCCCTGTCTGAGACTGAG--GCTGCCACTTCTGGCCTCAAGAATCAGAACGATG 2140
Db 29285 ACTCCAACCCTGTCTGAGACTGAGGCGCCTCCTCTTTCATCCTCGGGAATCAGAGGGATG 29344
Qy 2141 GGGACTCAGATTGCATGGGGGTGGGACCCAGGCCTGCAAGGCTTACGCGGAGGAAGAGGA 2200
Db 29345 GAGACTCACGTCAGCAGAGGGTGGGGCCCAACCCTGCCAGGATCAAGGAGAGGAAGAAGA 29404
Qy 2201 GGGAGGACTCAGGGGACCTTGGAATCCAGATCAGTGTGGACCTCGGCCCTGAGAGGTCCA 2260
Db 29405 GGGAGGACTCAGGGTACCTTTGAGTCCAGAACAATGGGGACCTTTGCCCTGGGAGGTCCA 29464
Qy 2261 GGGCACGGTGGCCACATATGGCCCATATTTCCTGCATCTTTGAGGTGAC----AGGACAG 2316
Db 29465 GTGCACAGTGGCCACCTGTAGCCCATGCTTGCTGCACCTTCTGGGTGACAAAGAGGAGAG 29524
Qy 2317 AGCTGTGGTCTGAGAAGTGGGGCCTCAGGTCAACAGAGGGAGGAGTTCCAGGATCCATAT 2376
Db 29525 GGCTGTGGTCAGAGCAGTGGTGACTCAGGTCAGCAGAGGGAGGAGTCCCAGCATCTGCAG 29584
Qy 2377 GGCCCAAGATGTGCCCCCTTCATGAGGACTGGGGATATCCCCGGCTCAGAAAGAAGGGAC 2436
Db 29585 GCCCCAATGTGTGCCCCATTCATGAAGATTGGGGATA-CCTTGGCTCAGAAAGAAGGGAC 29643
Qy 2437 TCCACACAGTCTGGCTGTCCCCTTTTAGTAGCTCTAGGGGGACCAGATCAGGGATGGCGG 2496
Db 29644 CCCACAGAGTCTGGCTGTCCCCTGATTTTTGCTCAGAGGGGACCAAATCAAGGATAGCCC 29703
Qy 2497 TATGTTCCATTCTCACTTGTACCACAGGCAGGAAGTTGGGGGGCCCTCAGGGAGATGGGG 2556
Db 29704 TATGTGCCAACCTCATTTGTGCCACAGGAAAGAAGTTGAAGAGCCCTCAGGGTGATGGGG 29763
Qy 2557 TCTTGGGGTAAAGGGGGGATGTCTACTCATGTCAGGGAATTGGGGGTTGAGGAAGCACAG 2616
Db 29764 TCTTGCAGTAAAGGGGAGCTATCTGCTCATCTCAGGGGGTTTCAGGTTGAGGAATGGCAG 29823
Qy 2617 GCGCTGGCAGGAATAAAGATGAGTGAGACAGACAAGGCTATTGGAATCCACACCCCAGAA 2676
Db 29824 GCCCCATCACG----ATGAAGAGTAACCCACAGGAGGCCATAGAAACACTCACCCCAGAA 29879
Qy 2677 CCAAAGGGGTCAGCCCTGGACACCTCA---------CCCAGGATGTGGCTTCTTTTTCAC 2727
Db 29880 CCAAAGGGGTCATACCTGGACACCCCATGTGGGGGTGACAGGATGTAGC-TCCATCTCAT 29938
Qy 2728 TCCTGTTTCCAGATCTGGGGCAGGTGAGGACCTCATTCTCAGAGGGTGACTCAGGTCAAC 2787
Db 29939 TCCTGTTTTCAGATCTCGGGGAGGTGAGGAACTTGTTCTCCGAGGATGACTCAGGTCAAC 29998
Qy 2788 GTAGGGACCCCCATCTGGTCTAAAGACAGAGCGGTCCCAGGATCTGCCATGCGTTCGGGT 2847
Db 29999 ACAGGGGCCCCCATCTGGTGGATAGACAGAGTGGTCCCAGGATCTGTCAGTAGTTCCGGT 30058
Qy 2848 GAGGAACATGAGGGAGGACTGAGGGTACCCCAGGACCAGAACACTGAGGGAGACTGCACA 2907
Db 30059 GAGGAACATGAGGGACGATTGAGGGCACCCTTGGGCCAGAACACAGATGAGGACCTCACG 30118
Qy 2908 GAAATCAGCCCTGCCCCCTGCTGTCACCCCAGAGAGCATGGGCTGGGCCGTCTGCCGAGGT 2967
Db 30119 GAAATCTGCCCTGCCCCCTGCTGTCACTCCAGAGAGCATGGGCAGGGCTGTCTGCTGCAGT 30178
Qy 2968 CCTTCC--GTTATCCTGGGATCATTGATGTCAGGGACGGGGAGGCCTTGGTCTGAGAAGG 3025
Db 30179 CCCCCCCACTTACCCTGGGATCATTGGTGTCAGGGATGGGGAGGTCTTTGTC-GAGGGGT 30237
Qy 3026 CTGCGCTCAGGTCAGTAGAGGGAGCGTCCCAGGCCCTGCCAGGAGTCAAGGTGAGGACCA 3085
Db 30238 CTGCACTCAGGTCAGTAGAGGGAGCGTCTTAGGCCCTGCCAGGAGACAAGGTAAGAACGA 30297
Qy 3086 AGCGGGCACCTCACCCAGGACACATTAATTCCAATGAATTTTGATATCTCTTGCTGCCCT 3145
Db 30298 AGCAGGTTCCTCACCCAGGACACATGAATTCCAATGCATTTCAGCATCTCTTCCTGTCCT 30357
Qy 3146 TCCCCA-AGGACCTAGGCACGTGTGGCCAGATGTTTGTCCCCTCCTGTCCTTCCATTCCT 3204
Db 30358 TCCCAAGAGGACCTGGGCACGTGTGGCCAGATGTGAGTCTCCTCATGTCCT---GTTCCC 30414
Qy 3205 TATCATGGATGTGAACTCTTGATTTGGATTTCTCAGACCAGCAAAAGGGCAGGATCCAGG 3264
Db 30415 TATCAGGGATGTGAGCTCTTAATCTGAGTTTCTCAGGCCAGCAAAAGGGTGGGATCCAGG 30474
Qy 3265 CCCTGCCAGGAAAAATATAAGGGCCCTGCGTGAGAACAGAGGGGGTCATCCACTGCATGA 3324
Db 30475 CCTTGCCAGGAGAAAGGTGAGGGCCCTGTGTGAGCACAGAGGGGACCATTCACCCCAAGA 30534
Qy 3325 GAGTGGGGATGTCACAGAGTCCAGCCCACCCTCCTGGTAGCACTGAGAAGCCAGGGCTGT 3384
Db 30535 GGGTGGAGACCTCACAGATTCCAGCCTACCCTCCTGTTAGCACTGGGGGCCTGAGGCTGT 30594
Qy 3385 GCTTGCGGTCTGCACCCTGAGGGCCCGTGGATTCCTCTTCCTGGAGCTCCAGGAACCAGG 3444
Db 30595 GCTTGCAGTCTGCACCCTGAGGGCCCATGCATTCCTCTTCCAGGAGCTCCAGGAAACAGA 30654
Qy 3445 CAGTGAGGCCTTGGTCTGAGACAGTATCCTCAGGTCACAGAGCAGAGGATGCACAG---- 3500
Db 30655 CACTGAGGCCTTGGTCTGAGGCCGTGCCCTCAGGTCACAGAGCAGAGGAGATGCAGACGT 30714
Qy 3501 GGTGTGCCAGCAGTGAATGTTTGCCCTGAATGCACACCAAGGGCCCCACCTGCCACAGGA 3560
Db 30715 CTAGTGCCAGCAGTGAACGTTTGCCTTGAATGCACACTAATGGCCCCCATCGCCCCAGAA 30774
Qy 3561 CACATAGGACTCCACAGAGTCTGGCCTCA-CCTCCCTACTGTCAGTCCTGTAGAATCGAC 3619
Db 30775 CATATGGGACTCCAGAGCACCTGGCCTCACCCTCTCTACTGTCAGTCCTGCAGAATCAGC 30834
Qy 3620 CTCTGCTGGCCGGCTGTACCCTGAGTACCCTCTCACTTCCTCCTTCAGGTTTTCAGGGGA 3679
Db 30835 CTCTGCTTGCTTGTGTACCCTGAGGTGCCCTCTCACTTTTTCCTTCAGGTTCTCAGGGGA 30894
Qy 3680 CAGGCCAACCCAGAGGAC----------AGGATTCCCTGGAGGCCACAGAGGAGCAC 3726
Db 30895 CAGGCTGACCAGGATCACCAGGAAGCTCCAGAGGATCCCCAGGAGGCCCTAGAGGAGCAC 30954
Qy 3727 C-AAGGAGAAGATCTGTAAGTAGGCCTTTGTTAGAGTCTCCAAGGTTCAGTTCTCAGCTG 3785
Db 30955 CAAAGGAGAAGATCTGTAAGTAAGCCTTTGTTAGAGCCTCCAAGGTTCAGTTTTTAGCTG 31014
Qy 3786 AGGCCTCTCACACACTCCCTCTCTCCCCAGGCCTGTGGGTCTTCATTGCCCAGCTCCTGC 3845
Db 31015 AGGCTTCTCACATGCTCCCTCTCTCTCCAGGCCAGTGGGTCTCCATTGCCCAGCTCCTGC 31074
Qy 3846 CCACACTCCTGCCTGCTGCCCTGACGAGAGTCATCATGTCTCTTGAGCAGAGGAGTCTGC 3905
Db 31075 CCACACTCCTGCCTGTTGCGGTGACCAGAGTCGTCATGTCTCTTGAGCAGAAGAGTCAGC 31134
Qy 3906 ACTGCAAGCCTGAGGAAGCCCTTGAGGCCCAACAAGAGGCCCTGGGCCTGGTGTGTGTGC 3965
Db 31135 ACTGCAAGCCTGAGGAAGGCCTTGACACCCAAGAAGAGGCCCTGGGCCTGGTGGGTGTGC 31194
Qy 3966 AGGCTGCCACC----------TCCTCCTCCTCTCCTCTGGTCCTGGGCA 4004
Db 31195 AGGCTGCCACTACTGAGGAGCAGGAGGCTGTGTCCTCCTCCTCTCCTCTGGTCCCAGGCA 31254
Qy 4005 CCCTGGAGGAGGTGCCCACTGCTGGGTCAACAGATCCTCCCCAGAGTCCTCAGGGAGCCT 4064
Db 31255 CCCTGGGGGAGGTGCCTGCTGCTGGGTCACCAGGTCCTCTCAAGAGTCCTCAGGGAGCCT 31314
Qy 4065 CCGCCTTTCCCACTACCATCAACTTCACTCGACAGAGGCAACCCAGTGAGGGTTCCAGCA 4124
Db 31315 CCGCCATCCCCACTGCCATCGATTTCACTCTATGGAGGCAATCCATTAAGGGCTCCAGCA 31374
Qy 4125 GCCGTGAAGAGGAGGGGCCAAGCACCTCTTGTATCCTGGAGTCCTTGTTCCGAGCAGTAA 4184
Db 31375 ACCAAGAAGAGGAGGGGCCAAGCACCTCCCCTGACCCAGAGTCTGTGTTCCGAGCAGCAC 31434

```
Qy    4185 TCACTAAGAAGGTGGCTGATTTGGTTGGTTTTCTGCTCCTCAAATATCGAGCCAGGGAGC 4244
Db   31435 TCAGTAAGAAGGTGGCTGACTTGATTCATTTTCTGCTCCTCAAGTATTAAGTCAAGGAGC 31494

Qy    4245 CAGTCACAAAGGCAGAAATGCTGGAGAGTGTCATCAAAAATTACAAGCACTGTTTTCCTG 4304
Db   31495 CGGTCACAAAGGCAGAAATGCTGGAGAGCGTCATCAAAAATTACAAGCGCTGCTTTCCTG 31554

Qy    4305 AGATCTTCGGCAAAGCCTCTGAGTCCTTGCAGCTGGTCTTTGGCATTGACGTGAAGGAAG 4364
Db   31555 AGATCTTCGGCAAAGCCTCCGAGTCCTTGCAGCTGGTCTTTGGCATTGACGTGAAGGAAG 31614

Qy    4365 CAGACCCCACCGGCCACTCCTATGTCCTTGTCACCTGCCTAGGTCTCTCCTATGATGGCC 4424
Db   31615 CGGACCCCACCAGCAACACCTACACCCTTGTCACCTGCCTGGG--ACTCCTATGATGGCC 31672

Qy    4425 TGCTGGGTGATAATCAGATCATGCCCAAGACAGGCTTCCTGATAATTGTCCTGGTCATGA 4484
Db   31673 TGCTGGTT--TAATCAGATCATGCCCAAGACGGGCCTCCTGATAATCGTCTTGGGCATGA 31730

Qy    4485 TTGCAATGGAGGGCGGCCATGCTCCTGAGGAGGAAATCTGGGAGGAGCTGAGTGTGATGG 4544
Db   31731 TTGCAATGGAGGGCAAATGCGTCCCTGAGGAGAAAATCTGGGAGGAGCTGAGTGTGATGA 31790

Qy    4545 AGGTGTATGATGGGAGGGAGCACAGTGCCTATGGGGAGCCCAGGAAGCTGCTCACCCAAG 4604
Db   31791 AGGTGTATGTTGGGAGGGAGCACAGTGTCTGTGGGGAGCCCAGGAAGCTGCTCACCCAAG 31850

Qy    4605 ATTTGGTGCAGGAAAAGTACCTGGAGTA-CGGCAGGTGCCGGACAGTGATCCCGCACGCT 4663
Db   31851 ATTTGGTGCAGGAAAACTACCTGGAGTACCGGCAGGTGCCCAGCAGTGATCCCATATGCT 31910

Qy    4664 ATGAGTTCCTGTGGGGTCCAAGGGCCCTCGCTGAAACCAGCTATGTGAAAGTCCTTGAGT 4723
Db   31911 ATGAGTTACTGTGGGGTCCAAGGGCACTCGCTG----------CTTGAAAGTACTGGAGC 31960

Qy    4724 ATGTGATCAAGGTCAGTGCAAGAGTTCGCTTTTTCTTCCCATCCCTGCGTGAAGCAGCTT 4783
Db   31961 ACGTGGTCAGGGTCAATGCAAGAGTTCTCATTTCCTACCCATCCCTGCGTGAAGCAGCTT 32020

Qy    4784 TGAGAGAGGAGGAAGAGGGAGTCTGAGCATGAGTTGCAGCCAAGGCCAGTGGGAGGGGGA 4843
Db   32021 TGAGAGAGGAGGAAGAGGGAGTCTGAGCATGAGCTGCAGCCAGGGCCACTGCGAGGGGGG 32080

Qy    4844 CTGGGCCAGTGCACCTTCCAGGGCCGCGTCCAGCAGCTTCCCCTGCC-TCGTGTGACATG 4902
Db   32081 CTGGGCCAGTGCACCTTCCAGGGCTCCGTCCAGTAGTTTCCCCTGCCTTAATGTGACATG 32140

Qy    4903 AGGCCCATTCTTCACTC--TGAAGAGAGCGGTCAGTGTTCTCAGTAGTAGGTTTCTGTTC 4960
Db   32141 AGGCCCATTCTTCTCTCTTTGAAGAGAGCAGTCAACATTCTTAGTAGTGGGTTTCTGTTC 32200

Qy    4961 TATTGGGTGACTTGGAGATTTATCTTTGTTCTCTTTTGGAATTGTTCAAATGTTTTTTTT 5020
Db   32201 TATTGGATGACTTTGAGATTTGTCTTTGTTTCCTTTTGGAATTGTTCAAATGTTTC-TTT 32259

Qy    5021 TAAGGGATGGTTGAATGAACTTCAGCATCCAAGTTTATGAATGACAGCAGTCACAC--AG 5078
Db   32260 TAATGGGTGGTTGAATGAACTTCAGCATTCAAATTTATGAATGACAGTAGTCACACATAG 32319

Qy    5079 TTCTGTGTATATAGTTTAAGGGTAAGAGTCTTGTGTTTTATTCAGATTGGGAAATCCATT 5138
Db   32320 TGCTGTTTATATAGTTTAGGAGTAAGAGTCTTGTTTTTTATTCAGATTGGGAAATCCATT 32379

Qy    5139 CTATTTTGTGAATTGGG--ATAATAACAGCAGTGGAATAAGTACTTA-GAAATGTGAAAA 5195
Db   32380 CCATTTTGTGAATTGGGACATAGTTACAGCAGTGGAATAAGTATTCATTTAGAAATGTGA 32439

Qy    5196 ATGAGCAGTAAAATAGATGAGATAAAGAACTAAAGAAATTAAGAGATAGTCAATTCTTGC 5255
Db   32440 ATGAGCAGTAAAACTGATGACA--------TAAAGAAATTAAAAGATATTTAATTCTTGC 32491

Qy    5256 CTTATACCTCAGTCTATTCTGTAAAA--TTTTTAAAGATATATGCATACCTGGATTTCCT 5313
Db   32492 CTTATA-CTCAGTCTATTCGGTAAAATTTTTTTAAAAAATGTGCATACCTGGATTTCCT 32550

Qy    5314 TGGCTTCTTTGAGAATGTAAGAGAAATTAAATCTGAATAAAGAATTCTTCCTGTTCACTG 5373
Db   32551 TGGCTTCTTTGAGAATGTAAGACAAATTAAATCTGAATAAATCATTCTCCCTGTTCACTG 32610

Qy    5374 GCTCTTTTCTTCTCCATGCACTGAGCATCTGCTTTTTGGAAGGCCCTGGGTTAGTAGTGG 5433
Db   32611 GCTCATTTATTCTCTATGCACTGAGCATTTGCTCTGTGGAAGGCCCTGGGTTAATAGTGG 32670

Qy    5434 AGATGCTAAGGTAAGCCAGACTCATACCCACCCATAGGGTCGTAGAGTCTAGGAGCTGCA 5493
Db   32671 AGATGCTAAGGTAAGCCAGACTCACCCCTACCCACAGGGTAGTAAAGTCTAGGAGCAGCA 32730

Qy    5494 GTCACGTAATCGAGGTGGCAAGATGTCCTCTAAAGATGTAGGGAAAAGTGAGAGAGGGGT 5553
Db   32731 GTCATATAATTAAGGTGGAGAGATGCCCTCTAAGATGTAGAGAAAAAGTAAGAAAGGGGT 32790

Qy    5554 GAGGGTGTGGGGCTCCGGGTGAGAGTGGTGGAGTGTCAATGCCCTGAGCTGGGGCATTTT 5613
Db   32791 GAGGGTGTGGGGCTCCAGGTGAGAGTGGTCCAGAGGTAAATGCCCTGTGTGGGGCCTTTT 32850

Qy    5614 GGGCTTTGGGAAACTGCAGTTCCTTCTGGGGGAGCTGATTGTAATGATCTTGGGTGG 5670
Db   32851 GGACTTTGGGAAACTGCAGTTTCTTCGGAGGGAGCTGATTCTAATGAATCGGCGGGG 32907
```

```
RESULT      7
HSU10687
LOCUS       HSU10687     11495 bp    DNA             PRI       23-JUN-1995
DEFINITION  Human MAGE-4a antigen (MAGE4a) gene, complete cds.
ACCESSION   U10687
VERSION     U10687.1  GI:533514
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 11495)
  AUTHORS   De Plaen,E., Arden,K., Traversari,C., Gaforio,J.J., Szikora,J.P.,
            De Smet,C., Brasseur,F., van der Bruggen,P., Lethe,B., Lurquin,C.
            Brasseur,R., Chomez,P., De Backer,O., Cavenee,W. and Boon,T.
  TITLE     Structure, chromosomal localization, and expression of 12 genes of
            the MAGE family
  JOURNAL   Immunogenetics 40 (5), 360-369 (1994)
  MEDLINE   95012457
REFERENCE   2  (bases 1 to 11495)
  AUTHORS   De Plaen,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JUN-1994) Etienne De Plaen, Ludwig Institute for
            Cancer Research, 74 Avenue Hippocrate, Brussels, 1200, Belgium
FEATURES             Location/Qualifiers
     source          1..11495
                     /organism="Homo sapiens"
                     /isolate="patient MZ2"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /sex="female"
                     /cell_type="lymphocyte"
                     /tissue_type="blood"
                     /dev_stage="adult"
     exon            450..510
                     /note="exon 1-1"
     exon            836..937
                     /note="exon 1-2"
     exon            1418..1556
                     /note="exon 1-3"
     exon            2108..2246
                     /note="exon 1-4"
     exon            2836..2928
                     /note="exon 1-5"
     exon            3508..3607
```

```
                  /note="exon 1-6"
     exon         4906. .4963
                  /note="exon 1-7"
     exon         5519. .5648
                  /note="exon 1-8"
     exon         9377. .9448
                  /number=2
     exon         9524. .11091
                  /number=3
     gene         9589. .10542
                  /gene="MAGE4a"
     CDS          9589. .10542
                  /gene="MAGE4a"
                  /codon_start=1
                  /product="MAGE-4a antigen"
                  /protein_id="AAA68871.1"
                  /db_xref="GI:533515"
                  /translation="MSSEQKSQHCKPEEGVEAQEEALGLVGAQAPTTEEQEAAVSSSS
                  PLVPGTLEEVPAAESAGPPQSPQGASALPTTISFTCWRQPNEGSSSQEEEGPSTSPDA
                  ESLFREALSNKVDELAHFLLRKYRAKELVTKAEMLERVIKNYKRCFPVIFGKASESLK
                  MIFGIDVKEVDPASNTYTLVTCLGLSYDGLLGNNQIFPKTGLLIIVLGTIAMEGDSAS
                  EEEIWEELGVMGVYDGREHTVYGEPRKLLTQDWVQENYLEYRQVPGSNPARYEFLWGP
                  RALAETSYVKVLEHVVRVNARVRIAYPSLREAALLEEEEGV"
     polyA_signal 11072. .11077
     polyA_site   11091
BASE COUNT     2530 a   3788 c   2924 g   2253 t
ORIGIN

  Query Match            46.8%;  Score 2655;  DB 10;  Length 11495;
  Best Local Similarity  73.6%;  Pred. No. 0;
  Matches 4231;  Conservative    0;  Mismatches 1100;  Indels  416;  Gaps   50;

Qy        21 CCTCCCCCTACCACCCCCAATCCCTCCCTTTACGCCACCCATCCAAACATCTTCACGCTC 80
             || | ||    || ||    | |||  ||| |          | | |  | | |  |
Db      5969 CCAGCACCCCTATCCTCCCCAAACCCCCACTACCCTTATGTCCTCATCCCCCACCCCAAC 6028

Qy        81 ACCCCCAGCCCAAGCCAGGCAGAATCCGGTTCCACCCCTGCTCTCAACCCAGGGAAGC-C 139
             ||| | | ||| | |||||  |||||    ||||    ||||| |||||||||||||| |
Db      6029 ACCACTATCCCCATCCAGGTTGAATCGCATTCCGTTTCTGCTTTCAACCCAGGGAAGCTC 6088

Qy       140 CAGGTGCCCAGATGTGACGCCACTGACTTGAGCATTAGTGGTTAGAGAGAAGCGAGGTTT 199
             |||||  ||  |||||||  ||||  ||||||| ||||  |||||||||||  || || ||
Db      6089 CAGGTTCCTGGATGTGATGCCAGTGACTTGTGCATTGGGGGTTAGAGAGACGCTAGCTTC 6148

Qy       200 TCGGTCTGAGGGGCGGCTTGAGATCGGTGGAGGGAAGCGGGCCCA-GCTCTGTAAGGAGG 258
             || ||||||  |||  ||||| ||| ||  ||||||||| || ||| ||||||| ||| ||
Db      6149 TCAGTCTGACAGGCAGCTTGGGATTGGCAGAGGGAAGCCGGTCCAGGCTCTGTGAGGTGG 6208

Qy       259 CAAGGTGACATGCTGAGGGAGGACT---GAGGACCCACTTACCCAGATAGAGGACCCCA 315
             ||  |||| | ||||||||||  | |    |||| ||      |||||||||||| ||||||
Db      6209 CATAGTGAGAAGCTGAGGGAGAAGTCGGGAGGCCCTCTCCACCCCAGATAGACGACCCCA 6268

Qy       316 AATAATCC---------CTTCATGCCAGTCCTGGACCATCTGGTGGTGGACTTCTCAGGC 366
             ||||||||          |   | |||||||||| ||| |   || |||||||| ||  |
Db      6269 AATAATCCGGCACCCCTCCTGCTTCCAGTCCTGGGCCACCCGTGGGCGGACTTCTGAGTC 6328

Qy       367 TGGGCCACCCCCAGCCCCCTTGCTGCTTAAACCACTGGGGACTCGAAGTCAGAGCTCCGT 426
             |||| |  ||| |    ||||  ||| ||| ||| || || | |||      ||||||||||  ||
Db      6329 TGGGACGCCCACCACCCCACTGCCGCTGAAGCCGCAGGGACTATGGAGTCAGAGCTTGGT 6388

Qy       427 GTGATCAGGGAAGGGCTGCTTAGGAGAGGGCAGCGTCCAGGCTCTGCCAGACATCATGCT 486
             ||||||||  | ||| |||  |   ||              |||||||||||| |||||    |
Db      6389 GTGATCAGTGCAGGACTGGTGGGGGT-----------AGGCTCTGCCAGGCATCAACGT 6436

Qy       487 CAGGATTCTCAAGGAGGGCTGAGGGTCCCTAAGACCCCACTCCCGTGACCCAACCCCCAC 546
             |||||  ||    ||||||||||  |||||     ||| |    | |  |||    ||||
Db      6437 CAGGACCCTAGGAGAGGGCTGAGTGTCCCCCA---CCCCCATTCCTATCCCCTACCCCTT 6493

Qy       547 TCCAATGCTCACTCCCGTGACCCAACCCCCTCTTCATTGTCATTCCAACCCCCACCCCAC 606
             ||| ||   |||||||                          |||     |        ||| |
Db      6494 TCCCATCTGCACTCCC--------------------------TACCCCATCTGTACCCCC 6527

Qy       607 ATCCCCCACCCCATCCCTCAACCCTGATGCCCATCCGCCCAGCCATTCCACCCTCACCCC 666
             || |||||||      |||  | ||         ||  ||  |||||        ||| | ||||
Db      6528 ATTCCCCACCTGTGCCCCTATCCTCCCCAACCCCCCAACCAGCCTCATACCCCCCTCCCC 6587

Qy       667 CACCCCCACCCCCACGCCCACTCCCACCCCCACCCAGGCAGGATCCG-GTTCCCGCCAGG 725
             ||||||  |||   ||   ||||        |        |     |   |  |||   |||||
Db      6588 CACCCCTACCTTCATCCCCATCAGTGCAGCATCCGGTTCCACCCCTGCTTTCAATCCAGG 6647

Qy       726 AAACATCCGGGTGCCCGGATGTGACGCCACTGACTTGCGCATTGTGGGGCAGAGAGAAGC 785
              ||   | |||||  ||||||||||  |||||||||||| | |||| |||   ||||||||
Db      6648 CAAGCCCTGGGTGGCCGGATGTGATGCCACTGACTTGTGAATTGAGGGTTAGAGAGAAGT 6707

Qy       786 GAGGTTTCCATTCTGAGGGACGGCGTAGAGTTCGGCCGAAGGAACCTGACCCAGGCTCTG 845
             ||| ||      ||||| ||  ||| | ||| ||||| || ||||  ||  ||||||||| ||
Db      6708 GAGTTTCTGGGTCTGAAGGGTGGC-TTGAGATCGGCAGAGGGAAGGTGGCCCAGGCTTTG 6766

Qy       846 TGAGGAGGCAAGGTGAGAGGCTGAGGGAGGACTGAGGACCCCGCCACTCCAAATAGAGAG 905
             ||| ||||||| ||||||  ||||||||||| | ||||  || |  |   ||  |||||| |
Db      6767 TGAAGAGGCAAAGTGAGACTCTGAGGGAGGATTCAGGAAACCCCTATCCCTGATAGAGGG 6826

Qy       906 CCCCAAATATTCCAGCCCCGCCCTTGCTGCCAGCCCTGGCCCACCCGCGGGAAGACGTCT 965
              ||||     |       |||                       | ||  || ||  |||
Db      6827 TCCCAGCCCTGGACTACCC-----------------------CGCGGAGGCTGACTTCT 6862

Qy       966 CAGCCTGGGCTGCCCCCAGACCCCTGCTCCAAAAGCCTTGAGAGACACCAGGTTCTTCTC 1025
             ||| ||||||||| |||    | ||  | ||                | |            | ||
Db      6863 CAGACTGGGCTGCTCCCCACCTCCGCCCCC-----TTCGCAACGCGTTTGTTTAAGCCAC 6917

Qy      1026 CCCAAGCTCTGGAATCAGAGGTTGCTGTGACCAGGGCAGGACTGGTTAGGAGAGGGCAGG 1085
                      |||||| |||||||||| ||||| ||||| ||| ||||||||||| |||| |
Db      6918 AGGGGACTCTGGAGTCAGAGGTTGGTGTGATCAGGGAAGGGCTGGTTAGGAGA-GGCATG 6976

Qy      1086 GCACAGGCTCTGCCAGGCATCAAGATCAGCACCCAAGAGGGAGGGCTGTGGGCCCCCAAG 1145
             || |||||  |||||||| ||||  ||||  ||| ||  | ||     ||| || |||||||||
Db      6977 GCCCAGGCCCTGCCAGGAATCAAAGTCAGAAACC-TGAGAGGGAACTGAGGTCCCCCAAG 7035

Qy      1146 ACTGCACTCCAATCCCCACTCCCACCCCATTCGCATTCCCATTCCCCACCCAACCCCCAT 1205
             |      | || || |||||||||||
Db      7036 ATCCTAGTCTAACCCCCACTCCCACAA------------------------------ 7062

Qy      1206 CTCCTCAGCTACACCTCCACCCCCATCCCTACTCCTACTCCGTCACCTGACCACCACCCT 1265
                               | || |   || |       || ||| |  ||    |      ||||
Db      7063 -----------------ATCCGCTGCCATTTCGCTGCTCCATTTCCCATTCCTTGCCCT 7104

Qy      1266 CCAGCCCCAGCACCAGCCCCAACCCTTCTGCCACCTCACCCTCACTGCCCCCAACCCCAC 1325
             ||| || || ||
Db      7105 CCACCCTCACCA------------------------------------------- 7116

Qy      1326 CCTCATCTCTCTCATGTGCCCCACTCCCATCGCCTCCCCCATTCTGGCAGAATCCGGTTT 1385
                                                          |||||||||| |||
Db      7117 -------------------------------------------GGCAGAATCCAGTTC 7131

Qy      1386 GCC-CCTGCTCTCAACCCAGGGAAGCCCTGGTAGGCCCGATGTGAAACCACTGACTTGAA 1444
              ||  |||||  |||| ||||||||  |||   |          | || ||        |
Db      7132 CCCTTCTGCTATCAATCCAGGGAAACCCCAGGCTTGGTGCTGGGATGTTTTT-------- 7183

Qy      1445 CCTCACAGATCTGAGAGAAGCCAGGTTCATTTAATGGTTCTGAGGGGCGGCTTGAGATCC 1504
                         || | |        |       ||  || ||||||||     ||||||||
Db      7184 ----------TGGGGGTCAGAGAATCAAGGGCATAGTCCTGAGGGGCCAGTTGAGATCG 7232

Qy      1505 ACTGAGGGGAGTGGTTTAGGCTCTGTGAGGAGGCAAGGTGAGATGCTGAGGGAGGACTG 1564
             ||||||||| ||      |  ||||||||  |||||||||||||||  |||||| |||||||
Db      7233 GCTGAGGGGAGCGGGCCCAAGCTCTGTGGCGAGGCAAGGTGAGACTCTGAGGAAGGACTG 7292

Qy      1565 AGGAGGCACACACCCCAGGTAGATGGCCCCAAAATGATCCAGTACCACCCCTGCTGCCAG 1624
             |||||| | ||||| || |||| ||  || ||||  ||||||    |     | ||||||||||
Db      7293 AGGAGGCCCCCACCCAAGATAGA-GGAACCCAAATAATCCAGCGCAGCTCCTGCTGCCAG 7351

Qy      1625 CCCTGGACCACCCGGCCAGGACAGATGTCTCAGCTGGACCACCCCCCGTCCCGTCCCACT 1684
             |||||||||||||   ||  ||| |||||     |  |  ||     |     ||||||
Db      7352 TCCTGGACCACCCGG---GGGAAGACTTCTCA----GGCTAGGCCATCCCAGCTCCCACT 7404

Qy      1685 GCCACTTAACCCACAGGGCAATCTGTAGTCATAGCTTATGTGACCGGGGCAGGGTTGGTC 1744
```


Db 7405 GCCACTAAAGCTACAGGGGACTCTAGAGTCAAGAGCTTGGTGTGCCCA---------- 7452
Qy 1745 AGGAGAGGCAGGGCCCAGGCATCAAGGTCCAGCATCCGCCCGGCATTAGGGTCAGGACCC 1804
Db 7453 -----AGGCAGGGCCCAGGC----------TCTGCCTGGCATCGGGGTCAGGACCT 7493
Qy 1805 TGGGAGGGAACTGAGGGTTCCCCACCCACACCTGTCTCCTCATCTCCACCGCCACCCCAC 1864
Db 7494 TGAGAGGGAACTGAGGGCGCTACACCCCCACCCCATCCGCATTC----------CAACAT 7543
Qy 1865 TCACATTCCCATACCTACCCCCTACCCCCAACCTCATCTTGTCAGAATCCCTGCTGTCAA 1924
Db 7544 GCCCAGCCCCATCCCCAACTCCGTTTTGCAGAATCCATTTTT----TCCCCTGCAGTCAA 7599
Qy 1925 CCCACGGAAGCCACGGGAATGGCGGCCAGGCACTCGGATCTTGACGTCCCCATCCAGGGT 1984
Db 7600 CCCCGGGAAGACCTGGGAATGGT---CAGGCACTCGGATCTTGACATCCACATCGAGGGC 7656
Qy 1985 CTGATGGAGGGAAGGGG-------CTTGAACAGGGCCTCAGGGGAGCAGAGGGAG----- 2032
Db 7657 TGAAGGAGGGAGAGGGTTTGGTATCATGAGCAGAGCCTCAGGGTAGCAGAGGGAGGACCC 7716
Qy 2033 -GGCCCTACTGCGAGATGAGGGAGGCCTCAGAGGACCCAGCACCC---------TAGGAC 2082
Db 7717 TGGCCCTCCTGGGAGATGAGGAAGGCCTCAGGAGACCCAGCACCCCAAGGCAGGGAGCCC 7776
Qy 2083 ACCGCACCCCTGTCTGAGACTGAG--GCTGCCACTTCTGGCCTCAAGAATCAGAACGATG 2140
Db 7777 ACCCCACCCCTGTCTGAGAATGAGGTGCCTCCTCTTTTAGCCTCAGGAATCCAAGGGATG 7836
Qy 2141 GGGACTCAGATTGCAT--GGGGGTGGGACCCAGGCCTGCAAGGCTTACGCGGAGGAAGAG 2198
Db 7837 GCAACTCAGGTCAGCAGAGGGGTGGGTTCCAAGCCCTTCCAGGATCAAGGAAAGGAAGAC 7896
Qy 2199 GAGGGAGGACTCAGGGGACCTTGGAATCCAGATCAGTGTGGACCTCGGCCCTGAGAGGTC 2258
Db 7897 GAGGGAGGATTCAGGGGGCCTTGCATTCCAGATCAGTGGAGACCTGGGCCCTGGGAGGTC 7956
Qy 2259 CAGGGCACGGTGGCCACATATGGCCCATATTTCCTGCATCTTTGAGGT----GACAGGAC 2314
Db 7957 CTGGGCAAGGTAGCCACCTGTAGCTCATACTTCCTGCATCTTCGAGGTCACAGAGAGGAG 8016
Qy 2315 AGAGCTGTGGTCTGAGAAGTGGGGCCTCAGGTCAACAGAGGGAGGAGTTCCAGGATCCAT 2374
Db 8017 AGGGCTATGGTCTGAGGGGTGGTACTTCAGGTCCGCAGAGGGAGGAGTCCCAGGATCTAC 8076
Qy 2375 ATGGCCCAAGATGTGCC-CCCTTCATGAGGACTGGGGATATCCCCGGCTCAGAAAGAAGG 2433
Db 8077 AGGACCCAAGGTGTGCCACACTTCACGAGGAATGGGGATACCTGTGGCTCAGAAAGACGG 8136
Qy 2434 GACTCCACACAGTCTGGCTGTCCCCTTTTAGTAGCTCTAGGGGGACCAGATCAGGGATGG 2493
Db 8137 GACCCCACAGAGTCTGGCTGTCCCCTGTTCTTAGCTCAGGGGGGACCAGAGGAGGGATGG 8196
Qy 2494 CGGTATGTTCCATTCTCACTTGTACCACAGGCAGGAAGTTGGGGGGCCCTCAGGGAGATG 2553
Db 8197 CCCTATGTGCCAATTTCACTTGTTCCACAGGCAGGAAGTTGGGGAACCTTCAGGGAGATG 8256
Qy 2554 GGGTCTTGGGGTAAAGGGGGGATGTCTACTCATGTCAGGGAATTGGGGGTTGAGGAAGCA 2613
Db 8257 AGGTTTTGGAGTAAAGGGGCAATGTTTGCTCATCTCAGGGGGTTGGGGGTTGAGGAAGGG 8316
Qy 2614 CAGGCGCTGGCAGGAATAAAGATGAGTGAGACAGACAAGGCTATTGGAATCCACACCCCA 2673
Db 8317 CAGGCCCTGTCAGGAGCAAACATGAGT-ACCCACAGGAGGCCATCAGAACCCTCACCCCA 8375
Qy 2674 GAACCAAAGGGGTCAGCCCTGGACACCTCACCCAGGATGTG--------GCTTCTTTTTC 2725
Db 8376 GAACCAAAGGGGTCAGCCCTGGGCACCCCACACAGGGGTGACAGGATGTGGCTCCTTCTC 8435
Qy 2726 ACTCCTGTTTCCAGATCTGGGGCAGGTGAGGACCTCATTCTCAGAGGGTGACTCAGGTCA 2785
Db 8436 ATTTCTGATTCCAGATCTCAGTGAGGTGAGGACCTTGTTCTCAGAGGGTGACTCAGGTCA 8495
Qy 2786 ACGTAGGGACCCCCATCTGGTCTAAAGACAGAGCGGTCCCAGGATCTGCCATGCGTTCGG 2845
Db 8496 CCACAGGGACCCCCATCTGGTCTACAGACACAGTGGTCCCAGGATCTGCCAAGAGTCCTG 8555
Qy 2846 GTGAGGAACATGAGGGAGGACTGAGGGTACCCCAGGACCAGAACACTGAGGGAGACTGCA 2905
Db 8556 GTGAGGAATGTGAGGGAGGATTGAGGGTACCACAGGGCCAGAACGCAGATGATGACCCCA 8615
Qy 2906 CAGAAATCAGCCCTGCCCCTGCTGTCACCCCAGAGAGCATGGGCTGGGCCGTCTGCCGAG 2965
Db 8616 CAGAAATCAGCCCTGCTCCTGTTGTCACCCCAGAGAGCATGGGCTTGGCTTTCTGCTGAG 8675
Qy 2966 GTCCTTCCGTTATCCTGGGATCATTGATGTCAGGGACGGGGAGGCCTTGGTCTGAGAAGG 3025
Db 8676 GTCCCTCTCTTATCCTGGGATCACTGGTGTCACGGAGTGGGAGGCCTTGGTCTGAGGGGG 8735
Qy 3026 CTGCGCTCAGGTCAGTAGAGGGAGCGTCCCAGGCCCTGCCAGGAGTCAAGGTGAGGACCA 3085
Db 8736 CTGCACCCAGGTCAGTAGAGGGAGGGTCCCAGGCTCTGCCAGGAGTTGAGGTGAGGACCA 8795
Qy 3086 AGCGGGCACCTCACCCAGGACACATTAATTCCAATGAATTTTGATATCTCTTGCTGCCCT 3145
Db 8796 AGCAGGCTCCGCATCCAGGACACATGGGTTCCAATGAATTTCGACATCTTTTGCTGTCGT 8855
Qy 3146 TC-CCCAAGGACCTAGGCACGTGTGGCCAGATGTTTGTCCCCTCCTGTCCTTCCATTCCT 3204
Db 8856 TCTTCGGAAGACCTAGGCACAGGTGGCCAGATGTGGGGTTTCTTAGGTCCT---GTTCCC 8912
Qy 3205 TATCATGGATGTGAACTCTTGATTTGGATTTCTCAGACCAGCAAAAGGGCAGGATCCAGG 3264
Db 8913 TCTCAGGCATGTGAGCTCTTGATCTGAGTTTCTCAGGCCAGCAAAAGAGTGGGATCCAGG 8972
Qy 3265 CCCTGCCAGGAAAAATATAAGGGCCCTGCGTGAGAACAGAGGGGGTCATCCACTGCATGA 3324
Db 8973 CCCTGCCTGGAGAAATGTGAGGGCCCTGAGTGAACACAGTGGGGATCATCCACTCCATGA 9032
Qy 3325 GAGTGGGGATGTCACAGAGTCCAGCCCACCCTCCTGGTAGCACTGAGAAGCCAGGGCTGT 3384
Db 9033 GAGTGGGGACCTCACAGAGTCCAGCCTACCCTCTTGATGGCACTGAGGGACCGGGGCTGT 9092
Qy 3385 GCTTGCGGTCTGCACCCTGAGGGCCCGTGGATTCCTCTTCCTGGAGCTCCAGGAACCAGG 3444
Db 9093 GCTTACAGTCTGCACCCTAAGGGCCCATGGATTCCTCTCCTAGGAGCTCCAGGAACAAGG 9152
Qy 3445 CAGTGAGGCCTTGGTCTGAGACAGTATCCTCAGGTCACAGAGCAGAGGATGCACAGGGTG 3504
Db 9153 CAGTGAGGCCTTGGTCTGAGACAGTGTCCTCAGGTTACAGAGCAGAGGATGCACAGGCTG 9212
Qy 3505 TGCCAGCAGTGAATGTTTGCCCTGAATGCACACCAAGGGCCCCACCTGCCACAGGACACA 3564
Db 9213 TGCCAGCAGTGAATGTTTGCCCTGAATGCACACCAAGGGCCCCACCTGCCACAAGACACA 9272
Qy 3565 TAGGACTCCACAGAGTCTGGCCTCACCTCCCTACTGTCAGTCCTGTAGAATCGACCTCTG 3624
Db 9273 TAGGACTCCAAAGAGTCTGGCCTCACCTCCCTACCATCAATCCTGCAGAATCGACCTCTG 9332
Qy 3625 CTGGCCGGCTGTACCCTGA-GTACCCTCTCACTTCCTCCTTCAGGTTTTCAGGGGACAGG 3683
Db 9333 CTGGCCGGCTATACCCTGAGGTGCTCTCTCACTTCCTCCTTCAGGTTCTGAGCAGACAGG 9392
Qy 3684 CCAACCCAGAGGACAGGATTCCCTGGAGGCCACAGAGGAGCACCAAGGAGAAGATCTGTA 3743
Db 9393 CCAACCG-GAGGACAGGATTCCCTGGAGGCCACAGAGGAGCACCAAGGAGAAGATCTGTA 9451
Qy 3744 AGTAGGCCTTTGTTAGAGTCTCCAAGGTTCAGTTCTCAGCTGAGGCCTCTCACACACTCC 3803
Db 9452 AGTAAGCCTTTGTTAGAGCCTCTAAGATTTGGTTCTCAGCTGAGGTCTCTCACATGCTCC 9511
Qy 3804 CTCTCTCCCCAGGCCTGTGGGTCTTCATTGCCCAGCTCCTGCCCACACTCCTGCCTGCTG 3863
Db 9512 CTCTCTCCGTAGGCCTGTGGGTCCCCATTGCCCAGCTTTTGCCTGCACTCTTGCCTGCTG 9571

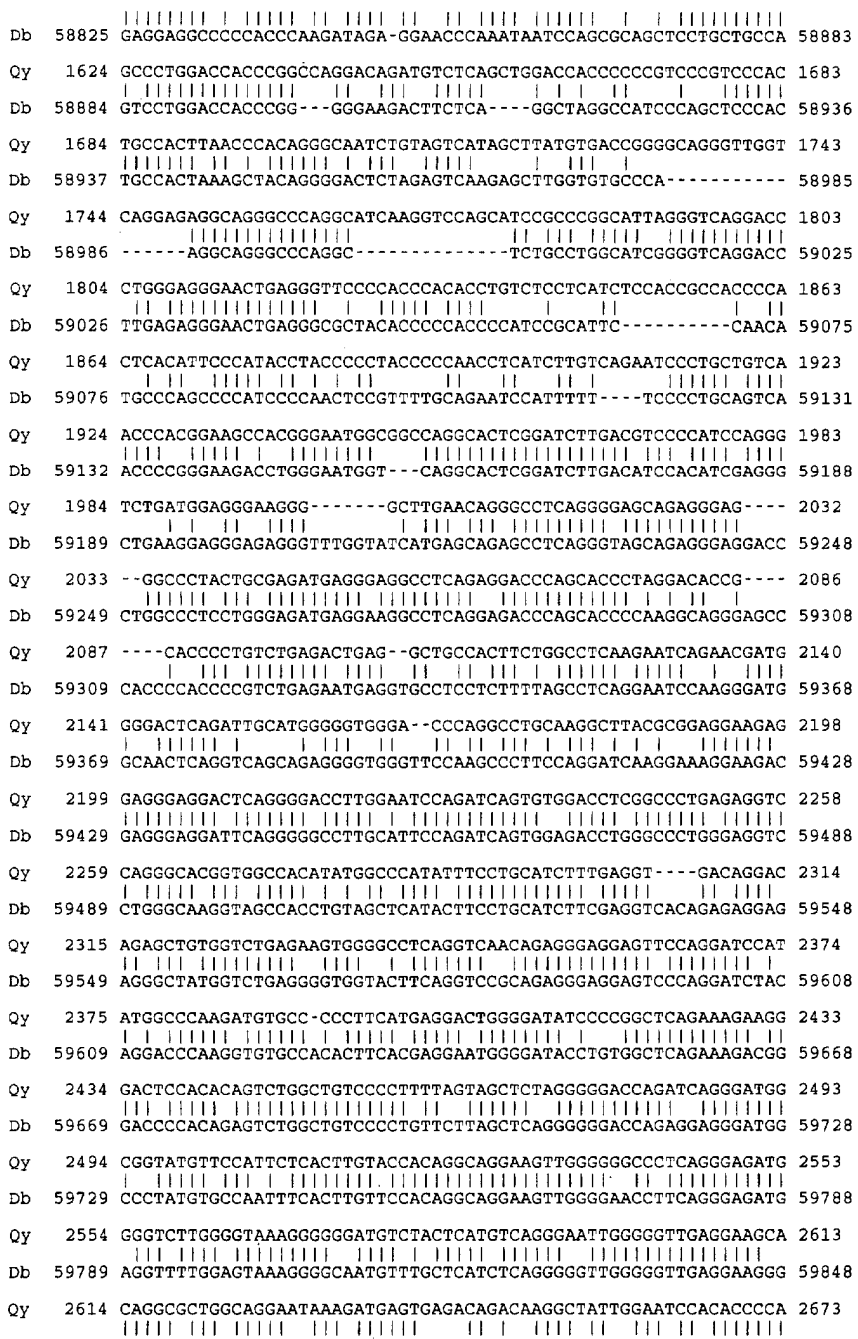
```
Qy   3864 CCCTGACGAGAGTCATCATGTCTCTTGAGCAGAGGAGTCTGCACTGCAAGCCTGAGGAAG 3923
Db   9572 CCCTGACCAGAGTCATCATGTCTTCTGAGCAGAAGAGTCAGCACTGCAAGCCTGAGGAAG 9631

Qy   3924 CCCTTGAGGCCCAACAAGAGGCCCTGGGCCTGGTGTGTGTGCAGGCTGCCAC-------- 3975
Db   9632 GCGTTGAGGCCCAAGAAGAGGCCCTGGGCCTGGTGGGTGCACAGGCTCCTACTACTGAGG 9691

Qy   3976 ----------------CTCCTCCTCCTCTCCTCTGGTCCTGGGCACCCTGGAGGAGGTGC 4019
Db   9692 AGCAGGAGGCTGCTGTCTCCTCCTCCTCTCCTCTGGTCCCTGGCACCCTGGAGGAAGTGC 9751

Qy   4020 CCACTGCTGGGTCAACAGATCCTCCCCAGAGTCCTCAGGGAGCCTCCGCCTTTCCCACTA 4079
Db   9752 CTGCTGCTGAGTCAGCAGGTCCTCCCCAGAGTCCTCAGGGAGCCTCTGCCTTACCCACTA 9811

Qy   4080 CCATCAACTTCACTCGACAGAGGCAACCCAGTGAGGGTTCCAGCAGCCGTGAAGAGGAGG 4139
Db   9812 CCATCAGCTTCACTTGCTGGAGGCAACCCAATGAGGGTTCCAGCAGCCAAGAAGAGGAGG 9871

Qy   4140 GGCCAAGCACCTCTTGTATCCTGGAGTCCTTGTTCCGAGCAGTAATCACTAAGAAGGTGG 4199
Db   9872 GGCCAAGCACCTCGCCTGACGCAGAGTCCTTGTTCCGAGAAGCACTCAGTAACAAGGTGG 9931

Qy   4200 CTGATTTGGTTGGTTTTCTGCTCCTCAAATATCGAGCCAGGGAGCCAGTCACAAAGGCAG 4259
Db   9932 ATGAGTTGGCTCATTTTCTGCTCCGCAAGTATCGAGCCAAGGAGCTGGTCACAAAGGCAG 9991

Qy   4260 AAATGCTGGAGAGTGTCATCAAAAATTACAAGCACTGTTTTCCTGAGATCTTCGGCAAAG 4319
Db   9992 AAATGCTGGAGAGAGTCATCAAAAATTACAAGCGCTGCTTTCCTGTGATCTTCGGCAAAG 10051

Qy   4320 CCTCTGAGTCCTTGCAGCTGGTCTTTGGCATTGACGTGAAGGAAGCAGACCCCACCGGCC 4379
Db   10052 CCTCCGAGTCCCTGAAGATGATCTTTGGCATTGACGTGAAGGAAGTGGACCCCGCCAGCA 10111

Qy   4380 ACTCCTATGTCCTTGTCACCTGCCTAGGTCTCTCCTATGATGGCCTGCTGGGTGATAATC 4439
Db   10112 ACACCTACACCCTTGTCACCTGCCTGGGCCTTTCCTATGATGGCCTGCTGGGTAATAATC 10171

Qy   4440 AGATCATGCCCAAGACAGGCTTCCTGATAATTGTCCTGGTCATGATTGCAATGGAGGGCG 4499
Db   10172 AGATCTTTCCCAAGACAGGCCTTCTGATAATCGTCCTGGGCACAATTGCAATGGAGGGCG 10231

Qy   4500 GCCATGCTCCTGAGGAGGAAATCTGGGAGGAGCTGAGTGTGATGGAGGTGTATGATGGGA 4559
Db   10232 ACAGCGCCTCTGAGGAGGAAATCTGGGAGGAGCTGGGTGTGATGGGGGTGTATGATGGGA 10291

Qy   4560 GGGAGCACAGTGCCTATGGGGAGCCCAGGAAGCTGCTCACCCAAGATTTGGTGCAGGAAA 4619
Db   10292 GGGAGCACACTGTCTATGGGGAGCCCAGGAAACTGCTCACCCAAGATTGGGTGCAGGAAA 10351

Qy   4620 AGTACCTGGAGTA-CGGCAGGTGCCGGACAGTGATCCCGCACGCTATGAGTTCCTGTGGG 4678
Db   10352 ACTACCTGGAGTACCGGCAGGTACCCGGCAGTAATCCTGCGCGCTATGAGTTCCTGTGGG 10411

Qy   4679 GTCCAAGGGCCCTCGCTGAAACCAGCTATGTGAAAGTCCTTGAGTATGTGATCAAGGTCA 4738
Db   10412 GTCCAAGGGCTCTGGCTGAAACCAGCTATGTGAAAGTCCTGGAGCATGTGGTCAGGGTCA 10471

Qy   4739 GTGCAAGAGTTCGCTTTTTCTTCCCATCCCTGCGTGAAGCAGCTTTGAGAGAGGAGGAAG 4798
Db   10472 ATGCAAGAGTTCGCATTGCCTACCCATCCCTGCGTGAAGCAGCTTTGTTAGAGGAGGAAG 10531

Qy   4799 AGGGAGTCTGAGCATGAGTTGCAGCCAAGGCCAGTGGGAGGG----GGACTGGGCCAGTG 4854
Db   10532 AGGGAGTCTGAGCATGAGTTGCAGCCAGGGCTGTGGGGAAGGGGCAGGGCTGGGCCAGTG 10591

Qy   4855 CACCTTCCAGGGCCGCGTCCAGCAGCTTCCCCTGCCTCGTGTGACATGAGGCCCATTCTT 4914
Db   10592 CATCTAACA--GCCCTGTGCAGCAGCTTCCCTTGCCTCGTGTAACATGAGGCCCATTCTT 10649
```

```
Qy   4915 CACTC----TGAAGAGAGCGGTCAGTGTTCTCAGTAGTAGGTTTCTGTTCTATTGGGTGA 4970
Db   10650 CACTCTGTTTGAAGAAAATAGTCAGTGTTCTTAGTAGTGGGTTTCTATTTTGTTGGATGA 10709

Qy   4971 CTTGGAGATTTATCTTTGTTCTCTTTTGGAATTGTTCAAATGTTTTTTTTTAAGGGATGG 5030
Db   10710 CTTGGAGATTTATCTCTGTTTCCTTTTACAATTGTTGAAATG-TTCCTTTTAATGGATGG 10768

Qy   5031 TTGAATGAACTTCAGCATCCAAGTTTATGAATGACAGCAGTCACACAGTTCTGTGTATAT 5090
Db   10769 TTGAATTAACTTCAGCATCCAAGTTTATGAATCGTAGTTAACGTATATTGCTGTTAATAT 10828

Qy   5091 AGTTTAAGGGTAAGAGTCTTGTGTTTTATTCAGATTGGGAAATCCATTCTATTTTGTGAA 5150
Db   10829 AGTTTAGGAGTAAGAGTCTTGTTTTTTATTCAGATTGGGAAATCCGTTCTATTTTGTGAA 10888

Qy   5151 TTGGG---ATAATAACAGCAGTGGAATAAGTACTTAGAAATGTGAAAAATGAGCAGTAAA 5207
Db   10889 TTTGGGACATAATAACAGCAGTGGAGTAAGTATTTAGAAGTGTG---AATTCACCGTGAA 10945

Qy   5208 ATAGATGAGATAAAGAACTAAAGAAATTAAGAGATAGTCAATTCTTGCCTTATACCTCAG 5267
Db   10946 ATAGGTGAGAT-----------AAATTAAAAGATACTTAATTCCCGCCTTATGCCTCAG 10993

Qy   5268 TCTATTCTGTAAAATTTTAAAGATATATGCATACCTGGATTTCCTTGGCTTCTTTGAGA 5327
Db   10994 TCTATTCTGTAAAATTTAAAAATATATATGCATACCTGGATTTCCTTGGCTTC---GTGA 11050

Qy   5328 ATGTAAGAGAAATTAAATCTGAATAAAGAATTCTTCCTGTTCACTGGCTCTTTTCTTCTC 5387
Db   11051 ATGTAAGAGAAATTAAATCTGAATAAATAATTCTTTCTGTTAACTGGCTCATTTCTTCTC 11110

Qy   5388 CATGCACTGAGCATCTGCTTTTTGGAAGGCCCTGGGTTAGTAGTGGAGATGCTAAGGTAA 5447
Db   11111 TATGCACTGAGCATCTGCTCTGTGGAAGGCCCAGGATTAGTAGTGGAGATACTAGGGTAA 11170

Qy   5448 GCCAGACTCATACCCACCCATAGGGTCGTAGAGTCTAGGAGCTGCAGTCACGTAATCGAG 5507
Db   11171 GCCAGACACACACCTACCGATAGGGTATTAAGAGTCTAGGAGCGCGGTCATATAATTAAG 11230

Qy   5508 GTGGCAAGATGTCCTCTAAAGATGTAGGGAAAAGTGAGAGAGGGGTGAGGGTGTGGGGCT 5567
Db   11231 GTGACAAGATGTCCTCTAAGATGTAGGGGAAAAGT----AACGAGTGTGGGTATGGGGCT 11286

Qy   5568 CCGGGTGAGAGTGGTGGAGTGTCAATGCCCTGAGCTGGGGCATTTTGGGCTTTGGGAAAC 5627
Db   11287 CCAGGTGAGAGTGGTCGGGTGTAAATTCCCTGTG-TGGGGCCTTTTGGGCTTTGGGAAAC 11345

Qy   5628 TGCAGTTCCTTCTGGGGGAGCTGATTGTAATGATCTTGGGTGGATCC 5674
Db   11346 TGCATTTTCTTCTGAGGGATCTGATTCTAATGAAGCTTGGTGGGTCC 11392
```

```
RESULT   8
U82696
LOCUS       U82696     158898 bp    DNA             PRI       01-JAN-1998
DEFINITION  Homo sapiens chromosome X map Xq28, complete sequence.
ACCESSION   U82696
VERSION     U82696.1  GI:2735037
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Hinzmann,B., Bauer,D., Drescher,B., Knop,A., Nordsiek,D.,
            Kioschis,P., Poustka,A., Rosenthal,A. and Gloeckner,G.
  TITLE     Location of GABA receptor gamma-4 subunit (GABRG4) on Xq28
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 158898)
  AUTHORS   Gloeckner,G., Rosenthal,A., Drescher,B., Schattevoy,R., Hinzmann,B.
            and Poustka,A.
  TITLE     Direct Submission
```

```
  JOURNAL   Submitted (19-DEC-1996) Genome Analysis, Institut for Molecular
            Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
FEATURES             Location/Qualifiers
     source          1. .158898
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /map="Xq28"
     misc_feature    1. .39138
                     /note="Region: cosmid IF064"
     repeat_region   1. .106
                     /rpt_family="L1PA11"
                     /evidence=not_experimental
     repeat_region   213. .282
                     /rpt_family="L1MB3"
                     /evidence=not_experimental
     exon            complement(1180. .1284)
                     /note="GRAIL, score = 43.000%, comment = marginal"
                     /evidence=not_experimental
     exon            complement(1683. .1767)
                     /note="MZEF, score = 60%"
                     /evidence=not_experimental
     exon            complement(2062. .2120)
                     /note="FEXHB, score = 11.76%"
                     /evidence=not_experimental
     repeat_region   complement(2637. .2710)
                     /rpt_family="L1ME2"
                     /evidence=not_experimental
     repeat_region   complement(2859. .3028)
                     /rpt_family="L1ME3A"
                     /evidence=not_experimental
     exon            complement(2911. .2978)
                     /note="MZEF, score = 79.7%"
                     /evidence=not_experimental
     exon            4250. .4431
                     /note="GRAIL, score = 52.000%, comment = good shadow"
                     /evidence=not_experimental
     repeat_region   complement(4639. .4866)
                     /rpt_family="L1MA2"
                     /evidence=not_experimental
     repeat_region   complement(4893. .5203)
                     /rpt_family="L1MA2"
                     /evidence=not_experimental
     exon            complement(4974. .5117)
                     /note="GRAIL, score = 61.000%, comment = good"
                     /evidence=not_experimental
     repeat_region   complement(5206. .5561)
                     /rpt_family="L1MA2"
                     /evidence=not_experimental
     repeat_region   complement(5562. .5602)
                     /rpt_family="L1MA9"
                     /evidence=not_experimental
     repeat_region   complement(5619. .5655)
                     /rpt_family="L1MA5"
                     /evidence=not_experimental
     repeat_region   complement(6084. .6254)
                     /rpt_family="MLT1D"
                     /evidence=not_experimental
     repeat_region   complement(6262. .6322)
                     /rpt_family="MLT1D"
                     /evidence=not_experimental
     repeat_region   complement(6384. .6549)
                     /rpt_family="MLT1D"
                     /evidence=not_experimental
     repeat_region   6664. .6771
                     /note="homology = 89.80%, score = 82, counts = 27"
                     /rpt_type=tandem
                     /rpt_unit=cctt
                     /evidence=not_experimental
     exon            6807. .6884
                     /note="FEXHB, score = 10.83%"
                     /evidence=not_experimental
     repeat_region   7337. .7454
                     /rpt_family="L1"
                     /evidence=not_experimental
     exon            7483. .7533
                     /note="FEXHB, score = 12.95%"
                     /evidence=not_experimental
     repeat_region   7950. .8010
                     /rpt_family="MSTA"
                     /evidence=not_experimental
     repeat_region   8147. .8238
                     /rpt_family="MSTA"
                     /evidence=not_experimental
     repeat_region   8245. .8286
                     /rpt_family="MSTA"
                     /evidence=not_experimental
     repeat_region   8287. .8361
                     /rpt_family="MSTC"
                     /evidence=not_experimental
     repeat_region   8878. .8921
                     /note="homology = 97.70%, score = 20, counts = 2"
                     /rpt_type=tandem
                     /rpt_unit=gattttaccttgtcaagtgcta
                     /evidence=not_experimental
     exon            complement(9471. .9526)
                     /note="MZEF, score = 85.6%"
                     /evidence=not_experimental
     exon            complement(9474. .9529)
                     /note="FEXHB, score = 12.02%"
                     /evidence=not_experimental
     repeat_region   9618. .9660
                     /rpt_family="MER3"
                     /evidence=not_experimental
     exon            10066. .10099
                     /note="GRAIL, score = 61.000%, comment = good"
                     /evidence=not_experimental
     exon            11110. .11190
                     /note="GRAIL, score = 76.000%, comment = excellent"
                     /evidence=not_experimental
     exon            11257. .11275
                     /note="FEXHB, score = 19.61%"
                     /evidence=not_experimental
     repeat_region   11360. .11419
                     /note="homology = 91.70%, score = 20, counts = 2"
                     /rpt_type=tandem
                     /rpt_unit=agggcaaatagctagaactattcatcaggc
                     /evidence=not_experimental
     exon            11613. .11637
                     /note="Xpound exon prediction, score = 68% (0%)"
                     /evidence=not_experimental
     exon            complement(12630. .12662)
                     /note="FEXHB, score = 13.07%"
                     /evidence=not_experimental
     repeat_region   13530. .13600
                     /rpt_family="MLT2B2"
                     /evidence=not_experimental
     repeat_region   13607. .13760
                     /rpt_family="MLT2B2"
                     /evidence=not_experimental
     exon            13820. .13947
                     /note="GRAIL, score = 88.000%, comment = excellent"
                     /evidence=not_experimental
     repeat_region   13924. .13977
                     /rpt_family="MLT2D"
                     /evidence=not_experimental
     repeat_region   13978. .14043
                     /note="homology = 75.80%, score = 23, counts = 6"
                     /rpt_type=tandem
                     /rpt_unit=ctctctccttt
                     /evidence=not_experimental
     repeat_region   14054. .14087
                     /note="homology = 94.10%, score = 28, counts = 17"
                     /rpt_type=tandem
                     /rpt_unit=tg
                     /evidence=not_experimental
```

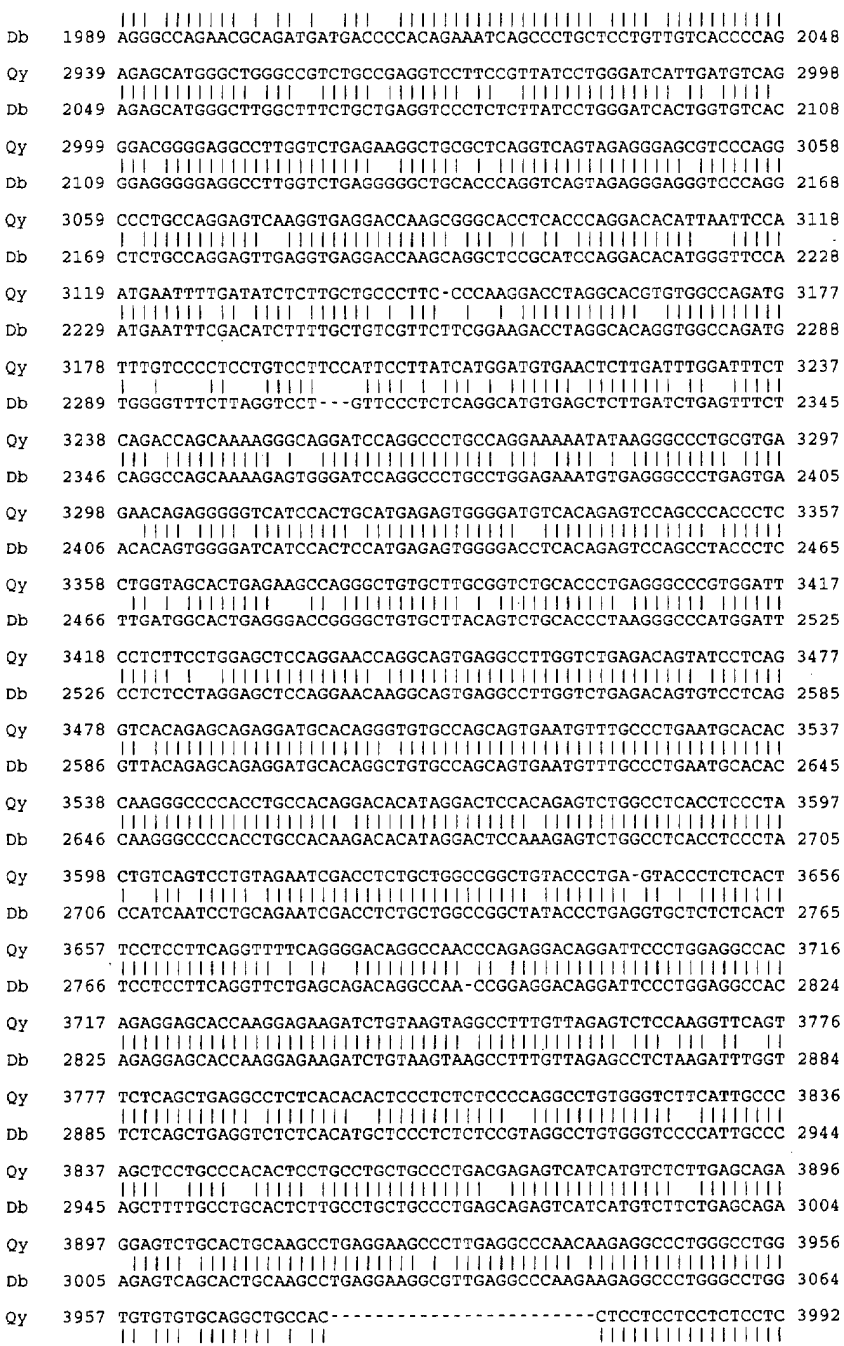
```
     exon            complement(15048. .15137)
                     /note="FEXHB, score = 12.14%"
                     /evidence=not_experimental
     exon            complement(15049. .15138)
                     /note="MZEF, score = 83.8%"
                     /evidence=not_experimental
     misc_feature    15192. .57359
                     /note="Region: Cosmid IG0799"
     exon            complement(15720. .15884)
                     /note="GRAIL, score = 53.000%, comment = good"
                     /evidence=not_experimental
     repeat_region   16505. .16631
                     /rpt_family="L1MA9"
                     /evidence=not_experimental
     repeat_region   16659. .16746
                     /rpt_family="MER4A"
                     /evidence=not_experimental
     repeat_region   16798. .16851
                     /rpt_family="MER4A"
                     /evidence=not_experimental
     repeat_region   complement(16851. .16910)
                     /rpt_family="THE1B"
                     /evidence=not_experimental
     repeat_region   complement(16920. .16972)
                     /rpt_family="THE1B"
                     /evidence=not_experimental
     exon            complement(16931. .16963)
                     /note="FEXHB, score = 12.61%"
                     /evidence=not_experimental
     repeat_region   complement(17023. .17178)
                     /rpt_family="THE1B"
                     /evidence=not_experimental
     repeat_region   complement(17252. .17279)
                     /rpt_family="THE1BR"
                     /evidence=not_experimental
     repeat_region   complement(17282. .17495)

  Query Match             46.8%;  Score 2653.2;  DB 11;  Length 158898;
  Best Local Similarity   73.6%;  Pred. No. 0;
  Matches 4230;  Conservative    0;  Mismatches 1103;  Indels  415;  Gaps   50;

Qy       21 CCTCCCCCTACCACCCCCAATCCCTCCCTTTACGCCACCCATCCAAACATCTTCACGCTC 80
Db    57500 CCAGCACCCCTATCCTCCCCAAACCCCCACTACCCTTATGTCCTCATCCCCCACCCCAAC 57559

Qy       81 ACCCCCAGCCCAAGCCAGGCAGAATCCGGTTCCACCCCTGCTCTCAACCCAGGGAAGC-C 139
Db    57560 ACCACTATCCCCATCCAGGTTGAATCGCATTCCGTTTCTGCTTTCAACCCAGGGAAGCTC 57619

Qy      140 CAGGTGCCCAGATGTGACGCCACTGACTTGAGCATTAGTGGTTAGAGAGAAGCGAGGTTT 199
Db    57620 CAGGTTCCTGGATGTGATGCCAGTGACTTGTGCATTGGGGGTTAGAGAGACGCTAGCTTC 57679

Qy      200 TCGGTCTGAGGGGCGGCTTGAGATCGGTGGAGGGAAGCGGGCCCA-GCTCTGTAAGGAGG 258
Db    57680 TCAGTCTGACAGGCAGCTTGGGATTGGCAGAGGGAAGCCGGTCCAGGCTCTGTGAGGTGG 57739

Qy      259 CAAGGTGACATGCTGAGGGAGGACT---GAGGACCCACTTACCCCAGATAGAGGACCCCA 315
Db    57740 CATAGTGAGAAGCTGAGGGAGAAGTCGGGAGGCCCTCTCCACCCCAGATAGACGACCCCA 57799

Qy      316 AATAATCC---------CTTCATGCCAGTCCTGGACCATCTGGTGGTGGACTTCTCAGGC 366
Db    57800 AATAATCCGGCACCCCTCCTGCTTCCAGTCCTGGGCCACCCGTGGGCGGACTTCTGAGTC 57859

Qy      367 TGGGCCACCCCCAGCCCCCTTGCTGCTTAAACCACTGGGGACTC-GAAGTCAGAGCTCCG 425
Db    57860 TGGGACGCCCACCACCCCACTGCCGCTGAAGCCGCAGGGGACTATGGAGTCAGAGCTTGG 57919

Qy      426 TGTGATCAGGGAAGGGCTGCTTAGGAGAGGGCAGCGTCCAGGCTCTGCCAGACATCATGC 485
Db    57920 TGTGATCAGTGCAGGACTGGTGGGGGT------------AGGCTCTGCCAGGCATCAACG 57967

Qy      486 TCAGGATTCTCAAGGAGGGCTGAGGGTCCCTAAGACCCCACTCCCGTGACCCAACCCCCA 545
Db    57968 TCAGGACCCTAGGAGAGGGCTGAGTGTCCCCCA---CCCCCATTCCTATCCCCTACCCCT 58024

Qy      546 CTCCAATGCTCACTCCCGTGACCCAACCCCCTCTTCATTGTCATTCCAACCCCCACCCCA 605
Db    58025 TTCCCATCTGCACTCCC-------------------------TACCCCATCTGTACCCC 58058

Qy      606 CATCCCCCACCCCATCCCTCAACCCTGATGCCCATCCGCCCAGCCATTCCACCCTCACCC 665
Db    58059 CATTCCCCACCTGTGCCCCTATCCTCCCCAACCCCCCAACCAGCCTCATACCCCCCTCCC 58118

Qy      666 CCACCCCCACCCCCACGCCCACTCCCACCCCCACCCAGGCAGGATCCG-GTTCCCGCCAG 724
Db    58119 CCACCCCTACCTTCATCCCCATCAGTGCAGCATCCGGTTCCACCCCTGCTTTCAATCCAG 58178

Qy      725 GAAACATCCGGGTGCCCGGATGTGACGCCACTGACTTGCGCATTGTGGGGCAGAGAGAAG 784
Db    58179 GCAAGCCCTGGGTGGCCGGATGTGATGCCACTGACTTGTGAATTGAGGGTTAGAGAGAAG 58238

Qy      785 CGAGGTTTCCATTCTGAGGGACGGCGTAGAGTTCGGCCGAAGGAACCTGACCCAGGCTCT 844
Db    58239 TGAGTTTCTGGGTCTGAAGGGTGGC-TTGAGATCGGCAGAGGGAAGGTGGCCCAGGCTTT 58297

Qy      845 GTGAGGAGGCAAGGTGAGAGGCTGAGGGAGGACTGAGGACCCCGCCACTCCAAATAGAGA 904
Db    58298 GTGAAGAGGCAAAGTGAGACTCTGAGGGAGGATTCAGGAAACCCCTATCCCTGATAGAGG 58357

Qy      905 GCCCCAAATATTCCAGCCCCGCCCTTGCTGCCAGCCCTGGCCCACCCGCGGGAAGACGTC 964
Db    58358 GTCCCAGCCCTGGACTACCC------------------------CGCGGAGGCTGACTTC 58393

Qy      965 TCAGCCTGGGCTGCCCCCAGACCCCTGCTCCAAAAGCCTTGAGAGACACCAGGTTCTTCT 1024
Db    58394 TCAGACTGGGCTGCTCCCCACCTCCGCCCCC-----TTCGCAACGCGTTTGTTTAAGCCA 58448

Qy     1025 CCCCAAGCTCTGGAATCAGAGGTTGCTGTGACCAGGGCAGGACTGGTTAGGAGAGGGCAG 1084
Db    58449 CAGGGGACTCTGGAGTCAGAGGTTGGTGTGATCAGGGAAGGGCTGGTTAGGAGAGGGCAT 58508

Qy     1085 GGCACAGGCTCTGCCAGGCATCAAGATCAGCACCCAAGAGGGAGGGCTGTGGGCCCCCAA 1144
Db    58509 GGCCCAGGCCCTGCCAGGAATCAAAGTCAGAAACC-TGAGAGGGAACTGAGGTCCCCCAA 58567

Qy     1145 GACTGCACTCCAATCCCCACTCCCACCCCATTCGCATTCCCATTCCCCACCCAACCCCCA 1204
Db    58568 GATCCTAGTCTAACCCCCACTCCCACA-------------------------------- 58594

Qy     1205 TCTCCTCAGCTACACCTCCACCCCCATCCCTACTCCTACTCCGTCACCTGACCACCACCC 1264
Db    58595 ------------------AATCCGCTGCCATTTCGCTGCTCCATTTCCCATTCCTTGCCC 58636

Qy     1265 TCCAGCCCCAGCACCAGCCCCAACCCTTCTGCCACCTCACCCTCACTGCCCCCAACCCCA 1324
Db    58637 TCCACCCTCACCA----------------------------------------------- 58649

Qy     1325 CCCTCATCTCTCTCATGTGCCCCACTCCCATCGCCTCCCCCATTCTGGCAGAATCCGGTT 1384
Db    58650 ----------------------------------------------GGCAGAATCCAGTT 58663

Qy     1385 TGCC-CCTGCTCTCAACCCAGGGAAGCCCTGGTAGGCCCGATGTGAAACCACTGACTTGA 1443
Db    58664 CCCCTTCTGCTATCAATCCAGGGAAACCCCAGGCTTGGTGCTGGGATGTTTTT------- 58716

Qy     1444 ACCTCACAGATCTGAGAGAAGCCAGGTTCATTTAATGGTTCTGAGGGGCGGCTTGAGATC 1503
Db    58717 ------------TGGGGGTCAGAGAATCAAGGGCATAGTCCTGAGGGGCCAGTTGAGATC 58764

Qy     1504 CACTGAGGGGAGTGGTTTTAGGCTCTGTGAGGAGGCAAGGTGAGATGCTGAGGGAGGACT 1563
Db    58765 GGCTGAGGGGAGCGGGCCCAAGCTCTGTGGCGAGGCAAGGTGAGACTCTGAGGAAGGACT 58824

Qy     1564 GAGGAGGCACACACCCCAGGTAGATGGCCCCAAAATGATCCAGTACCACCCCTGCTGCCA 1623
```


Db 58825 GAGGAGGCCCCCACCCAAGATAGA-GGAACCCAAATAATCCAGCGCAGCTCCTGCTGCCA 58883
Qy 1624 GCCCTGGACCACCCGGCCAGGACAGATGTCTCAGCTGGACCACCCCCCGTCCCGTCCCAC 1683
Db 58884 GTCCTGGACCACCCGG---GGGAAGACTTCTCA----GGCTAGGCCATCCCAGCTCCCAC 58936
Qy 1684 TGCCACTTAACCCACAGGGCAATCTGTAGTCATAGCTTATGTGACCGGGGCAGGGTTGGT 1743
Db 58937 TGCCACTAAAGCTACAGGGGACTCTAGAGTCAAGAGCTTGGTGTGCCCA---------- 58985
Qy 1744 CAGGAGAGGCAGGGCCCAGGCATCAAGGTCCAGCATCCGCCCGGCATTAGGGTCAGGACC 1803
Db 58986 ------AGGCAGGGCCCAGGC--------------TCTGCCTGGCATCGGGGTCAGGACC 59025
Qy 1804 CTGGGAGGGAACTGAGGGTTCCCCACCCACACCTGTCTCCTCATCTCCACCGCCACCCCA 1863
Db 59026 TTGAGAGGGAACTGAGGGCGCTACACCCCCACCCCATCCGCATTC----------CAACA 59075
Qy 1864 CTCACATTCCCATACCTACCCCCTACCCCCAACCTCATCTTGTCAGAATCCCTGCTGTCA 1923
Db 59076 TGCCCAGCCCCATCCCCAACTCCGTTTTGCAGAATCCATTTTT----TCCCCTGCAGTCA 59131
Qy 1924 ACCCACGGAAGCCACGGGAATGGCGGCCAGGCACTCGGATCTTGACGTCCCCATCCAGGG 1983
Db 59132 ACCCCGGGAAGACCTGGGAATGGT---CAGGCACTCGGATCTTGACATCCACATCGAGGG 59188
Qy 1984 TCTGATGGAGGGAAGGG-------GCTTGAACAGGGCCTCAGGGGAGCAGAGGGAG---- 2032
Db 59189 CTGAAGGAGGGAGAGGGTTTGGTATCATGAGCAGAGCCTCAGGGTAGCAGAGGGAGGACC 59248
Qy 2033 --GGCCCTACTGCGAGATGAGGGAGGCCTCAGAGGACCCAGCACCCTAGGACACCG---- 2086
Db 59249 CTGGCCCTCCTGGGAGATGAGGAAGGCCTCAGGAGACCCAGCACCCCAAGGCAGGGAGCC 59308
Qy 2087 ----CACCCCTGTCTGAGACTGAG--GCTGCCACTTCTGGCCTCAAGAATCAGAACGATG 2140
Db 59309 CACCCCACCCCGTCTGAGAATGAGGTGCCTCCTCTTTTAGCCTCAGGAATCCAAGGGATG 59368
Qy 2141 GGGACTCAGATTGCATGGGGGTGGGA--CCCAGGCCTGCAAGGCTTACGCGGAGGAAGAG 2198
Db 59369 GCAACTCAGGTCAGCAGAGGGGTGGGTTCCAAGCCCTTCCAGGATCAAGGAAAGGAAGAC 59428
Qy 2199 GAGGGAGGACTCAGGGGACCTTGGAATCCAGATCAGTGTGGACCTCGGCCCTGAGAGGTC 2258
Db 59429 GAGGGAGGATTCAGGGGGCCTTGCATTCCAGATCAGTGGAGACCTGGGCCCTGGGAGGTC 59488
Qy 2259 CAGGGCACGGTGGCCACATATGGCCCATATTTCCTGCATCTTTGAGGT----GACAGGAC 2314
Db 59489 CTGGGCAAGGTAGCCACCTGTAGCTCATACTTCCTGCATCTTCGAGGTCACAGAGAGGAG 59548
Qy 2315 AGAGCTGTGGTCTGAGAAGTGGGGCCTCAGGTCAACAGAGGGAGGAGTTCCAGGATCCAT 2374
Db 59549 AGGGCTATGGTCTGAGGGGTGGTACTTCAGGTCCGCAGAGGGAGGAGTCCCAGGATCTAC 59608
Qy 2375 ATGGCCCAAGATGTGCC-CCCTTCATGAGGACTGGGGATATCCCCGGCTCAGAAAGAAGG 2433
Db 59609 AGGACCCAAGGTGTGCCACACTTCACGAGGAATGGGGATACCTGTGGCTCAGAAAGACGG 59668
Qy 2434 GACTCCACACAGTCTGGCTGTCCCCTTTTAGTAGCTCTAGGGGGACCAGATCAGGGATGG 2493
Db 59669 GACCCCACAGAGTCTGGCTGTCCCCTGTTCTTAGCTCAGGGGGGACCAGAGGAGGGATGG 59728
Qy 2494 CGGTATGTTCCATTCTCACTTGTACCACAGGCAGGAAGTTGGGGGGCCCTCAGGGAGATG 2553
Db 59729 CCCTATGTGCCAATTTCACTTGTTCCACAGGCAGGAAGTTGGGGAACCTTCAGGGAGATG 59788
Qy 2554 GGGTCTTGGGGTAAAGGGGGGATGTCTACTCATGTCAGGGAATTGGGGGTTGAGGAAGCA 2613
Db 59789 AGGTTTTGGAGTAAAGGGGCAATGTTTGCTCATCTCAGGGGGTTGGGGGTTGAGGAAGGG 59848
Qy 2614 CAGGCGCTGGCAGGAATAAAGATGAGTGAGACAGACAAGGCTATTGGAATCCACACCCCA 2673
Db 59849 CAGGCCCTGTCAGGAGCAAACATGAGTACCCCACAGGAGGCCATCAGAACCCTCACCCCA 59908
Qy 2674 GAACCAAAGGGGTCAGCCCTGGACACCTCACCCAGGATGTG--------GCTTCTTTTTC 2725
Db 59909 GAACCAAAGGGGTCAGCCCTGGGCACCCCACACAGGGGTGACAGGATGTGGCTCCTTCTC 59968
Qy 2726 ACTCCTGTTTCCAGATCTGGGGCAGGTGAGGACCTCATTCTCAGAGGGTGACTCAGGTCA 2785
Db 59969 ATTTCTGATTCCAGATCTCAGTGAGGTGAGGACCTTGTTCTCAGAGGGTGACTCAGGTCA 60028
Qy 2786 ACGTAGGGACCCCCATCTGGTCTAAAGACAGAGCGGTCCCAGGATCTGCCATGCGTTCGG 2845
Db 60029 CCACAGGGACCCCCATCTGGTCTACAGACACAGTGGTCCCAGGATCTGCCAAGAGTCCTG 60088
Qy 2846 GTGAGGAACATGAGGGAGGACTGAGGGTACCCCAGGACCAGAACACTGAGGGAGACTGCA 2905
Db 60089 GTGAGGAATGTGAGGGAGGATTGAGGGTACCACAGGGCCAGAACGCAGATGATGACCCCA 60148
Qy 2906 CAGAAATCAGCCCTGCCCCTGCTGTCACCCCAGAGAGCATGGGCTGGGCCGTCTGCCGAG 2965
Db 60149 CAGAAATCAGCCCTGCTCCTGTTGTCACCCCAGAGAGCATGGGCTTGGCTTTCTGCTGAG 60208
Qy 2966 GTCCTTCCGTTATCCTGGGATCATTGATGTCAGGGACGGGGAGGCCTTGGTCTGAGAAGG 3025
Db 60209 GTCCCTCTCTTATCCTGGGATCACTGGTGTCACGGAGTGGGAGGCCTTGGTCTGAGGGGG 60268
Qy 3026 CTGCGCTCAGGTCAGTAGAGGGAGCGTCCCAGGCCCTGCCAGGAGTCAAGGTGAGGACCA 3085
Db 60269 CTGCACCCAGGTCAGTAGAGGGAGGGTCCCAGGCTCTGCCAGGAGTTGAGGTGAGGACCA 60328
Qy 3086 AGCGGGCACCTCACCCAGGACACATTAATTCCAATGAATTTTGATATCTCTTGCTGCCCT 3145
Db 60329 AGCAGGCTCCGCATCCAGGACACATGGGTTCCAATGAATTTCGACATCTTTTGCTGTCGT 60388
Qy 3146 TC-CCCAAGGACCTAGGCACGTGTGGCCAGATGTTTGTCCCCTCCTGTCCTTCCATTCCT 3204
Db 60389 TCTTCGGAAGACCTAGGCACAGGTGGCCAGATGTGGGGTTTCTTAGGTCCT---GTTCCC 60445
Qy 3205 TATCATGGATGTGAACTCTTGATTTGGATTTCTCAGACCAGCAAAAGGGCAGGATCCAGG 3264
Db 60446 TCTCAGGCATGTGAGCTCTTGATCTGAGTTTCTCAGGCCAGCAAAAGAGTGGGATCCAGG 60505
Qy 3265 CCCTGCCAGGAAAAATATAAGGGCCCTGCGTGAGAACAGAGGGGGTCATCCACTGCATGA 3324
Db 60506 CCCTGCCTGGAGAAATGTGAGGGCCCTGAGTGAACACAGTGGGGATCATCCACTCCATGA 60565
Qy 3325 GAGTGGGGATGTCACAGAGTCCAGCCCACCCTCCTGGTAGCACTGAGAAGCCAGGGCTGT 3384
Db 60566 GAGTGGGGACCTCACAGAGTCCAGCCTACCCTCTTGATGGCACTGAGGGACCGGGGCTGT 60625
Qy 3385 GCTTGCGGTCTGCACCCTGAGGGCCCGTGGATTCCTCTTCCTGGAGCTCCAGGAACCAGG 3444
Db 60626 GCTTACAGTCTGCACCCTAAGGGCCCATGGATTCCTCTCCTAGGAGCTCCAGGAACAAGG 60685
Qy 3445 CAGTGAGGCCTTGGTCTGAGACAGTATCCTCAGGTCACAGAGCAGAGGATGCACAGGGTG 3504
Db 60686 CAGTGAGGCCTTGGTCTGAGACAGTGTCCTCAGGTTACAGAGCAGAGGATGCACAGGCTG 60745
Qy 3505 TGCCAGCAGTGAATGTTTGCCCTGAATGCACACCAAGGGCCCCACCTGCCACAGGACACA 3564
Db 60746 TGCCAGCAGTGAATGTTTGCCCTGAATGCACACCAAGGGCCCCACCTGCCACAAGACACA 60805
Qy 3565 TAGGACTCCACAGAGTCTGGCCTCACCTCCCTACTGTCAGTCCTGTAGAATCGACCTCTG 3624
Db 60806 TAGGACTCCAAAGAGTCTGGCCTCACCTCCCTACCATCAATCCTGCAGAATCGACCTCTG 60865
Qy 3625 CTGGCCGGCTGTACCCTGA-GTACCCTCTCACTTCCTCCTTCAGGTTTTCAGGGGACAGG 3683
Db 60866 CTGGCCGGCTATACCCTGAGGTGCTCTCTCACTTCCTCCTTCAGGTTCTGAGCAGACAGG 60925
Qy 3684 CCAACCCAGAGGACAGGATTCCCTGGAGGCCACAGAGGAGCACCAAGGAGAAGATCTGTA 3743
Db 60926 CCAACCG-GAGGACAGGATTCCCTGGAGGCCACAGAGGAGCACCAAGGAGAAGATCTGTA 60984

```
Qy   3744 AGTAGGCCTTTGTTAGAGTCTCCAAGGTTCAGTTCTCAGCTGAGGCCTCTCACACACTCC 3803
Db  60985 AGTAAGCCTTTGTTAGAGCCTCTAAGATTTGGTTCTCAGCTGAGGTCTCTCACATGCTCC 61044

Qy   3804 CTCTCTCCCCAGGCCTGTGGGTCTTCATTGCCCAGCTCCTGCCCACACTCCTGCCTGCTG 3863
Db  61045 CTCTCTCCGTAGGCCTGTGGGTCCCCATTGCCCAGCTTTGCCTGCACTCTTGCCTGCTG 61104

Qy   3864 CCCTGACGAGAGTCATCATGTCTCTTGAGCAGAGGAGTCTGCACTGCAAGCCTGAGGAAG 3923
Db  61105 CCCTGACCAGAGTCATCATGTCTTCTGAGCAGAAGAGTCAGCACTGCAAGCCTGAGGAAG 61164

Qy   3924 CCCTTGAGGCCCAACAAGAGGCCCTGGGCCTGGTGTGTGTGCAGGCTGCCAC-------- 3975
Db  61165 GCGTTGAGGCCCAAGAAGAGGCCCTGGGCCTGGTGGGTGCACAGGCTCCTACTACTGAGG 61224

Qy   3976 ----------------CTCCTCCTCCTCTCCTCTGGTCCTGGGCACCCTGGAGGAGGTGC 4019
Db  61225 AGCAGGAGGCTGCTGTCTCCTCCTCCTCTCCTCTGGTCCCTGGCACCCTGGAGGAAGTGC 61284

Qy   4020 CCACTGCTGGGTCAACAGATCCTCCCCAGAGTCCTCAGGGAGCCTCCGCCTTTCCCACTA 4079
Db  61285 CTGCTGCTGAGTCAGCAGGTCCTCCCCAGAGTCCTCAGGGAGCCTCTGCCTTACCCACTA 61344

Qy   4080 CCATCAACTTCACTCGACAGAGGCAACCCAGTGAGGGTTCCAGCAGCCGTGAAGAGGAGG 4139
Db  61345 CCATCAGCTTCACTTGCTGGAGGCAACCCAATGAGGGTTCCAGCAGCCAAGAAGAGGAGG 61404

Qy   4140 GGCCAAGCACCTCTTGTATCCTGGAGTCCTTGTTCCGAGCAGTAATCACTAAGAAGGTGG 4199
Db  61405 GGCCAAGCACCTCGCCTGACGCAGAGTCCTTGTTCCGAGAAGCACTCAGTAACAAGGTGG 61464

Qy   4200 CTGATTTGGTTGGTTTTCTGCTCCTCAAATATCGAGCCAGGGAGCCAGTCACAAAGGCAG 4259
Db  61465 ATGAGTTGGCTCATTTTCTGCTCCGCAAGTATCGAGCCAAGGAGCTGGTCACAAAGGCAG 61524

Qy   4260 AAATGCTGGAGAGTGTCATCAAAAATTACAAGCACTGTTTTCCTGAGATCTTCGGCAAAG 4319
Db  61525 AAATGCTGGAGAGAGTCATCAAAAATTACAAGCGCTGCTTTCCTGTGATCTTCGGCAAAG 61584

Qy   4320 CCTCTGAGTCCTTGCAGCTGGTCTTTGGCATTGACGTGAAGGAAGCAGACCCCACCGGCC 4379
Db  61585 CCTCCGAGTCCCTGAAGATGATCTTTGGCATTGACGTGAAGGAAGTGGACCCCGCCAGCA 61644

Qy   4380 ACTCCTATGTCCTTGTCACCTGCCTAGGTCTCTCCTATGATGGCCTGCTGGGTGATAATC 4439
Db  61645 ACACCTACACCCTTGTCACCTGCCTGGGCCTTTCCTATGATGGCCTGCTGGGTAATAATC 61704

Qy   4440 AGATCATGCCCAAGACAGGCTTCCTGATAATTGTCCTGGTCATGATTGCAATGGAGGGCG 4499
Db  61705 AGATCTTTCCCAAGACAGGCCTTCTGATAATCGTCCTGGGCACAATTGCAATGGAGGGCG 61764

Qy   4500 GCCATGCTCCTGAGGAGGAAATCTGGGAGGAGCTGAGTGTGATGGAGGTGTATGATGGGA 4559
Db  61765 ACAGCGCCTCTGAGGAGGAAATCTGGGAGGAGCTGGGTGTGATGGGGGTGTATGATGGGA 61824

Qy   4560 GGGAGCACAGTGCCTATGGGGAGCCCAGGAAGCTGCTCACCCAAGATTTGGTGCAGGAAA 4619
Db  61825 GGGAGCACACTGTCTATGGGGAGCCCAGGAAACTGCTCACCCAAGATTGGGTGCAGGAAA 61884

Qy   4620 AGTACCTGGAGTA-CGGCAGGTGCCGGACAGTGATCCCGCACGCTATGAGTTCCTGTGGG 4678
Db  61885 ACTACCTGGAGTACCGGCAGGTACCCGGCAGTAATCCTGCGCGCTATGAGTTCCTGTGGG 61944

Qy   4679 GTCCAAGGGCCCTCGCTGAAACCAGCTATGTGAAAGTCCTTGAGTATGTGATCAAGGTCA 4738
Db  61945 GTCCAAGGGCTCTGGCTGAAACCAGCTATGTGAAAGTCCTGGAGCATGTGGTCAGGGTCA 62004

Qy   4739 GTGCAAGAGTTCGCTTTTTCTTCCCATCCCTGCGTGAAGCAGCTTTGAGAGAGGAGGAAG 4798
Db  62005 ATGCAAGAGTTCGCATTGCCTACCCATCCCTGCGTGAAGCAGCTTTGTTAGAGGAGGAAG 62064

Qy   4799 AGGGAGTCTGAGCATGAGTTGCAGCCAAGGCCAGTGGGAGGG----GGACTGGGCCAGTG 4854
Db  62065 AGGGAGTCTGAGCATGAGTTGCAGCCAGGGCTGTGGGGAAGGGGCAGGGCTGGGCCAGTG 62124

Qy   4855 CACCTTCCAGGGCCGCGTCCAGCAGCTTCCCCTGCCTCGTGTGACATGAGGCCCATTCTT 4914
Db  62125 CATCTAACA--GCCCTGTGCAGCAGCTTCCCTTGCCTCGTGTAACATGAGGCCCATTCTT 62182

Qy   4915 CACTC----TGAAGAGAGCGGTCAGTGTTCTCAGTAGTAGGTTTCTGTTCTATTGGGTGA 4970
Db  62183 CACTCTGTTTGAAGAAAATAGTCAGTGTTCTTAGTAGTGGGTTTCTATTTTGTTGGATGA 62242

Qy   4971 CTTGGAGATTTATCTTTGTTCTCTTTTGGAATTGTTCAAATGTTTTTTTTTAAGGGATGG 5030
Db  62243 CTTGGAGATTTATCTCTGTTTCCTTTTACAATTGTTGAAATG-TTCCTTTTAATGGATGG 62301

Qy   5031 TTGAATGAACTTCAGCATCCAAGTTTATGAATGACAGCAGTCACACAGTTCTGTGTATAT 5090
Db  62302 TTGAATTAACTTCAGCATCCAAGTTTATGAATCGTAGTTAACGTATATTGCTGTTAATAT 62361

Qy   5091 AGTTTAAGGGTAAGAGTCTTGTGTTTTATTCAGATTGGGAAATCCATTCTATTTTGTGAA 5150
Db  62362 AGTTTAGGAGTAAGAGTCTTGTTTTTTATTCAGATTGGGAAATCCGTTCTATTTTGTGAA 62421

Qy   5151 TTGGG---ATAATAACAGCAGTGGAATAAGTACTTAGAAATGTGAAAAATGAGCAGTAAA 5207
Db  62422 TTTGGGACATAATAACAGCAGTGGAGTAAGTATTTAGAAGTGTG---AATTCACCGTGAA 62478

Qy   5208 ATAGATGAGATAAAGAACTAAAGAAATTAAGAGATAGTCAATTCTTGCCTTATACCTCAG 5267
Db  62479 ATAGGTGAGAT------------AAATTAAAAGATACTTAATTCCCGCCTTATGCCTCAG 62526

Qy   5268 TCTATTCTGTAAAATTT-TTAAAGATATATGCATACCTGGATTTCCTTGGCTTCTTTGAG 5325
Db  62527 TCTATTCTGTAAAATTTAAAAAATATATATGCATACCTGGATTTCCTTGGCTTC---GTG 62583

Qy   5327 AATGTAAGAGAAATTAAATCTGAATAAAGAATTCTTCCTGTTCACTGGCTCTTTTCTTCT 5386
Db  62584 AATGTAAGAGAAATTAAATCTGAATAAATAATTCTTTCTGTTAACTGGCTCATTTCTTCT 62643

Qy   5387 CCATGCACTGAGCATCTGCTTTTTGGAAGGCCCTGGGTTAGTAGTGGAGATGCTAAGGTA 5446
Db  62644 CTATGCACTGAGCATCTGCTCTGTGGAAGGCCCAGGATTAGTAGTGGAGATACTAGGGTA 62703

Qy   5447 AGCCAGACTCATACCCACCCATAGGGTCGTAGAGTCTAGGAGCTGCAGTCACGTAATCGA 5506
Db  62704 AGCCAGACACACACCTACCGATAGGGTATTAAGAGTCTAGGAGCGCGGTCATATAATTAA 62763

Qy   5507 GGTGGCAAGATGTCCTCTAAAGATGTAGGGAAAAGTGAGAGAGGGGTGAGGGTGTGGGGC 5566
Db  62764 GGTGACAAGATGTCCTCTAAGATGTAGGGGAAAAGT----AACGAGTGTGGGTATGGGGC 62819

Qy   5567 TCCGGGTGAGAGTGGTGGAGTGTCAATGCCCTGAGCTGGGGCATTTTGGGCTTTGGGAAA 5626
Db  62820 TCCAGGTGAGAGTGGTCGGGTGTAAATTCCCTGTG-TGGGGCCTTTTGGGCTTTGGGAAA 62878

Qy   5627 CTGCAGTTCCTTCTGGGGGAGCTGATTGTAATGATCTTGGGTGGATCC 5674
Db  62879 CTCCATTTTCTTCTGAGGGATCTGATTCTAATGAAGCTTGGTGGGTCC 62926
```

```
RESULT   9
HSU10688
LOCUS       HSU10688     4895 bp    DNA             PRI       23-JUN-1995
DEFINITION  Human MAGE-4b antigen (MAGE4b) gene, complete cds.
ACCESSION   U10688
VERSION     U10688.1  GI:533516
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4895)
```

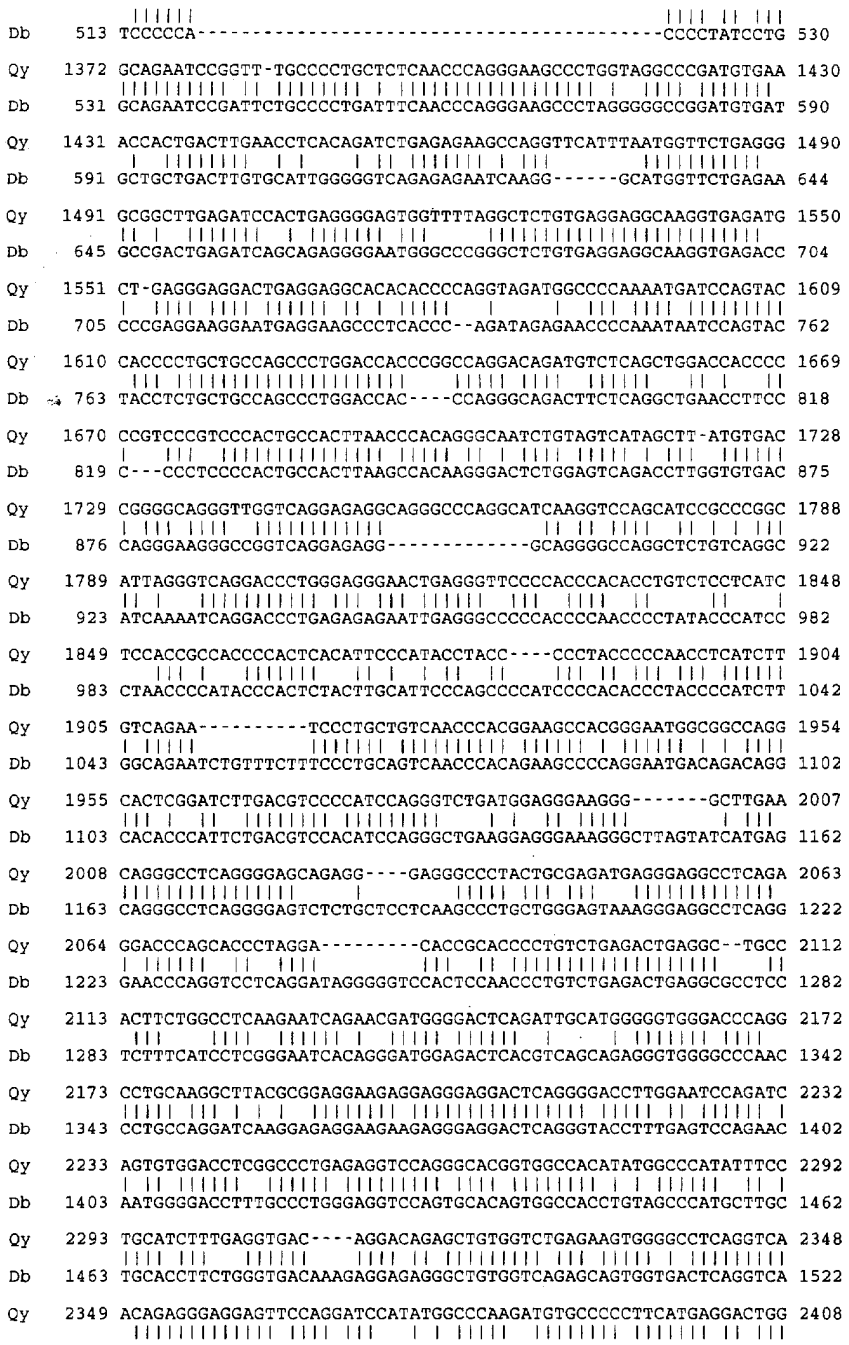
```
  AUTHORS   De Plaen,E., Arden,K., Traversari,C., Gaforio,J.J., Szikora,J.P.,
            De Smet,C., Brasseur,F., van der Bruggen,P., Lethe,B., Lurquin,C.
            Brasseur,R., Chomez,P., De Backer,O., Cavenee,W. and Boon,T.
  TITLE     Structure, chromosomal localization, and expression of 12 genes of
            the MAGE family
  JOURNAL   Immunogenetics 40 (5), 360-369 (1994)
  MEDLINE   95012457
REFERENCE   2  (bases 1 to 4895)
  AUTHORS   De Plaen,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JUN-1994) Etienne De Plaen, Ludwig Institute for
            Cancer Research, 74 Avenue Hippocrate, Brussels, 1200, Belgium
FEATURES             Location/Qualifiers
     source          1..4895
                     /organism="Homo sapiens"
                     /isolate="patient MZ2"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /sex="female"
                     /cell_type="lymphocyte"
                     /tissue_type="blood"
                     /dev_stage="adult"
     exon            2777..2848
                     /number=2
     exon            2924..4491
                     /number=3
     gene            2989..3942
                     /gene="MAGE4b"
     CDS             2989..3942
                     /gene="MAGE4b"
                     /codon_start=1
                     /product="MAGE-4b antigen"
                     /protein_id="AAA68872.1"
                     /db_xref="GI:533517"
                     /translation="MSSEQKSQHCKPEEGVEAQEEALGLVGAQAPTTEEQEAAVSSSS
                     PLVPGTLEEVPAAESAGPPQSPQGASALPTTISFTCWRQPNEGSSSQEEEGPSTSPDA
                     ESLFREALSNKVDELAHFLLRKYRAKELVTKAEMLERVIKNYKRCFPVIFGKASESLK
                     MIFGIDVKEVDPTSNTYTLVTCLGLSYDGLLGNNQIFPKTGLLIIVLGTIAMEGDSAS
                     EEEIWEELGVMGVYDGREHTVYGEPRKLLTQDWVQENYLEYRQVPGSNPARYEFLWGP
                     RALAETSYVKVLEHVVRVNARVRIAYPSLREAALLEEEEGV"
     polyA_signal    4472..4477
BASE COUNT     1134 a   1235 c   1433 g   1093 t
ORIGIN

  Query Match             44.3%;  Score 2513.6;  DB 10;  Length 4895;
  Best Local Similarity   79.9%;  Pred. No. 0;
  Matches 3415;  Conservative    0;  Mismatches  694;  Indels  165;  Gaps   32;

Qy       1478 ATGGTTCTGAGGGGCGGCTTGAGATCCACTGAGGGGAGTGGTTTTAGGCTCTGTGAGGAG 1537
Db        607 ATAGTCCTGAGGGGCCAGTTGAGATCGGCTGAGGGGAGCGGGCCCAAGCTCTGTGGCGAG 666

Qy       1538 GCAAGGTGAGATGCTGAGGGAGGACTGAGGAGGCACACACCCCAGGTAGATGGCCCCAAA 1597
Db        667 GCAAGGTGAGACTCTGAGGAAGGACTGAGGAGGCCCCCACCCAAGATAGA-GGAACCCAA 725

Qy       1598 ATGATCCAGTACCACCCCTGCTGCCAGCCCTGGACCACCCGGCCAGGACAGATGTCTCAG 1657
Db        726 ATAATCCAGCCCACGTCCTGCTGCCAGTCCTGGACCACCCGG---GGGAAGACTTCTCA- 781

Qy       1658 CTGGACCACCCCCCGTCCCGTCCCACTGCCACTTAACCCACAGGGCAATCTGTAGTCATA 1717
Db        782 ---GGCTAGGCCATCCCAGCTCCCACTGCCACTAAAGCTACAGGGGACTCTAGAGTCA-- 836

Qy       1718 GCTTATGTGACCGGGGCAGGGTTGGTCAGGAGAGGCAGGGCCCAGGCATCAAGGTCCAGC 1777
Db        837 ------------------------------AGAGCTTGGTGTGCCCAAGGCAGGGCCAGG 866

Qy       1778 ATCCGCCCGGCATTAGGGTCAGGACCCTGGGAGGGAACTGAGGGTTCCCCACCCACACCT 1837
Db        867 CTCTGCCTGGCATCGGGGTCAGGACCTTGAGAGGGAACTGAGGGCGCTACACCCCCACCC 926

Qy       1838 GTCTCCTCATCTCCACCGCCACCCCACTCACATTCCCATACCTACCCCCTACCCCCAACC 1897
Db        927 CATCCGCATTC----------CAACATGCCCAGCCCCATCCCCAACTCCGTTTTGCAGAA 976

Qy       1898 TCATCTTGTCAGAATCCCTGCTGTCAACCCACGGAAGCCACGGGAATGGCGGCCAGGCAC 1957
Db        977 TCCATTTTT----TCCCCTGCAGTCAACCCCGGGAAGACCTGGGAATGGT---CAGGCAC 1029

Qy       1958 TCGGATCTTGACGTCCCCATCCAGGGTCTGATGGAGGGAAGGG-------GCTTGAACAG 2010
Db       1030 TCGGATCTTGACATCCACATCGAGGGCTGAAGGAGGGAGAGAGTTTGGTATCATGAGCAG 1089

Qy       2011 GGCCTCAGGGGAGCAGAGGGAG------GGCCCTACTGCGAGATGAGGGAGGCCTCAGAG 2064
Db       1090 AGCCTCAGGGTAGCAGAGGGAGGACCCTGGCCCTCCTGGGAGATGAGGAAGGCCTCAGGA 1149

Qy       2065 GACCCAGCACCC---------TAGGACACCGCACCCCTGTCTGAGACTGAG--GCTGCCA 2113
Db       1150 GACCCAGCACCCCAAGGCAGGGAGCCCACCCCACCCCTGTCTGAGAATGAGGTGCCTCCT 1209

Qy       2114 CTTCTGGCCTCAAGAATCAGAACGATGGGGACTCAGATTGCATGGGGGTGGGA--CCCAG 2171
Db       1210 CCTTTAGCCTCAGGAATCCAAGGGATGGCAACTCAGGTCAGCAGAGGGGTGGGTTCCAAG 1269

Qy       2172 GCCTGCAAGGCTTACGCGGAGGAAGAGGAGGGAGGACTCAGGGGACCTTGGAATCCAGAT 2231
Db       1270 CCCTTCCAGGATCAAGGAAAGGAAGACGAGGGAGGATTCAGGGGGCCTTGCATTCCAGAT 1329

Qy       2232 CAGTGTGGACCTCGGCCCTGAGAGGTCCAGGGCACGGTGGCCACATATGGCCCATATTTC 2291
Db       1330 CAGTGGAGACCTGGGCCCTGGGAGGTCCTGGGCAAGGTAGCCACCTGTAGCTCATACTTC 1389

Qy       2292 CTGCATCTTTGAGGT----GACAGGACAGAGCTGTGGTCTGAGAAGTGGGGCCTCAGGTC 2347
Db       1390 CTGCATCTTCGAGGTCACAGAGAGGAGAGGGCTATGGTCTGAGGGGTGGTACTTCAGGTC 1449

Qy       2348 AACAGAGGGAGGAGTTCCAGGATCCATATGGCCCAAGATGTGCC-CCCTTCATGAGGACT 2406
Db       1450 CGCAGAGGGAGGAGTCCCAGGATCTACAGGACCCAAGGTGTGCCACACTTCACGAGGAAT 1509

Qy       2407 GGGGATATCCCCGGCTCAGAAAGAAGGGACTCCACACAGTCTGGCTGTCCCCTTTTAGTA 2466
Db       1510 GGGGATACCTGTGGCTCAGAAAGACGGGACCCCACAGAGTCTGGCTGTCCCCTGTTCTTA 1569

Qy       2467 GCTCTAGGGGGACCAGATCAGGGATGGCGGTATGTTCCATTCTCACTTGTACCACAGGCA 2526
Db       1570 GCTCAGGGGGGACCAGAGGAGGGATGGCCCTATGTGCCAATTTCACTTGTTCCACAGGCA 1629

Qy       2527 GGAAGTTGGGGGGCCCTCAGGGAGATGGGGTCTTGGGGTAAAGGGGGGATGTCTACTCAT 2586
Db       1630 GGAAGTTGGGGAACCTTCAGGGAGATGAGGTTTTGGAGTAAAGGGGCAATGTTTGCTCAT 1689

Qy       2587 GTCAGGGAATTGGGGGTTGAGGAAGCACAGGCGCTGGCAGGAATAAAGATGAGTGAGACA 2646
Db       1690 CTCAGGGGGTTGGGGGTTGAGGAAGGGCAGGCCCTGTCAGGAGCAAACATGAGT-ACCCA 1748

Qy       2647 GACAAGGCTATTGGAATCCACACCCCAGAACCAAAGGGGTCAGCCCTGGACACCTCACCC 2706
Db       1749 CAGGAGGCCATCAGAACCCTCACCCCAGAACCAAAGGGGTCAGCCCTGGGCACCCCACAC 1808

Qy       2707 AG--------GATGTGGCTTCTTTTTCACTCCTGTTTCCAGATCTGGGGCAGGTGAGGAC 2758
Db       1809 AGGGGTGACAGGATGTGGCTCCTTCTCATTTCTGATTCCAGATCTCAGTGAGGTGAGGAC 1868

Qy       2759 CTCATTCTCAGAGGGTGACTCAGGTCAACGTAGGGACCCCCATCTGGTCTAAAGACAGAG 2818
Db       1869 CTTGTTCTCAGAGGGTGACTCAGGTCACCACAGGGACCCCCATCTGGTCTACAGACACAG 1928

Qy       2819 CGGTCCCAGGATCTGCCATGCGTTCGGGTGAGGAACATGAGGGAGGACTGAGGGTACCCC 2878
Db       1929 TGGTCCCAGGATCTGCCAAGAGTCCTGGTGAGGAATGTGAGGGAGGATTGAGGGTACCAC 1988

Qy       2879 AGGACCAGAACACTGAGGGAGACTGCACAGAAATCAGCCCTGCCCCTGCTGTCACCCCAG 2938
```

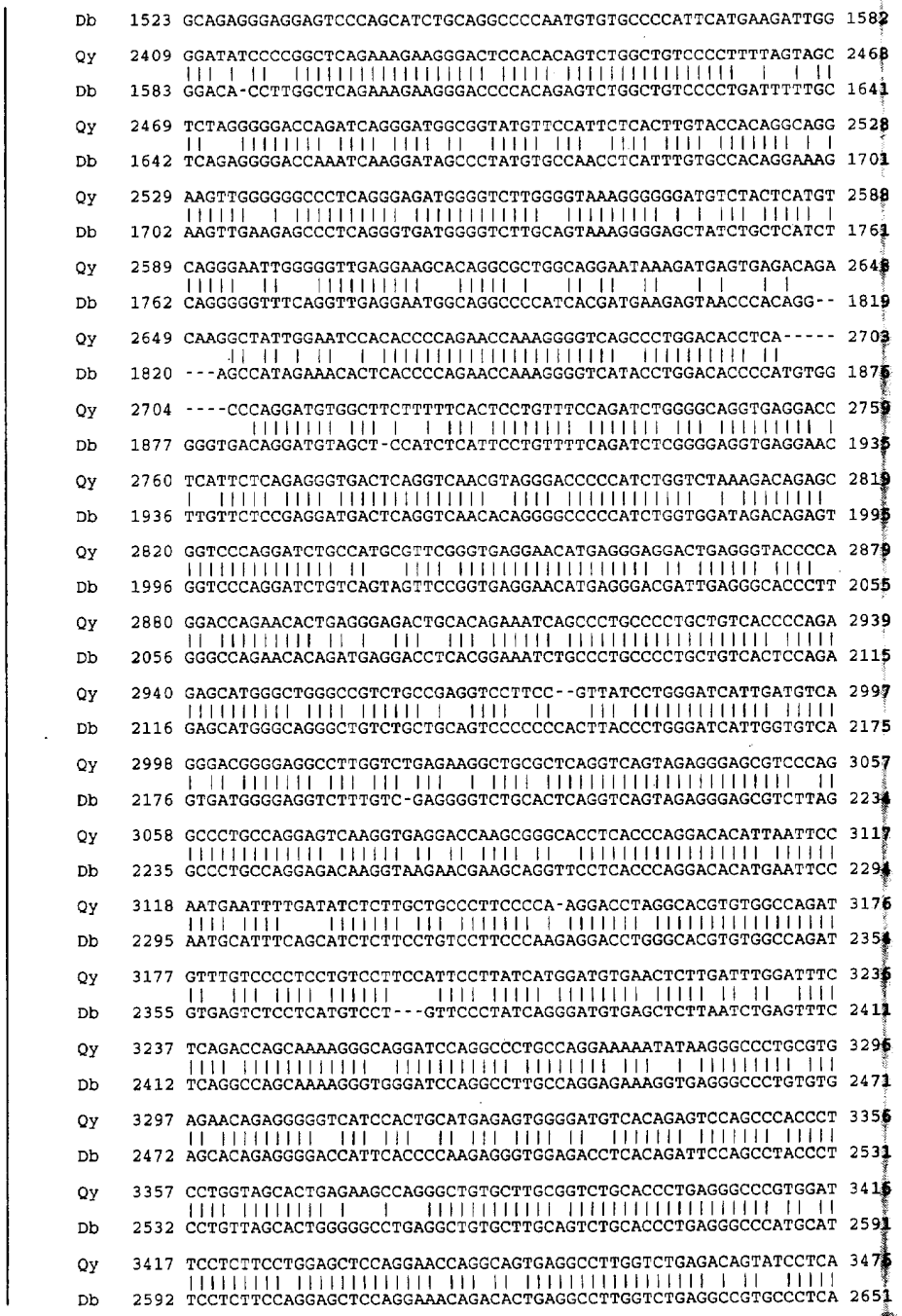

Db 1989 AGGGCCAGAACGCAGATGATGACCCCACAGAAATCAGCCCTGCTCCTGTTGTCACCCCAG 2048
Qy 2939 AGAGCATGGGCTGGGCCGTCTGCCGAGGTCCTTCCGTTATCCTGGGATCATTGATGTCAG 2998
Db 2049 AGAGCATGGGCTTGGCTTTCTGCTGAGGTCCCTCTCTTATCCTGGGATCACTGGTGTCAC 2108
Qy 2999 GGACGGGGAGGCCTTGGTCTGAGAAGGCTGCGCTCAGGTCAGTAGAGGGAGCGTCCCAGG 3058
Db 2109 GGAGGGGGAGGCCTTGGTCTGAGGGGCTGCACCCAGGTCAGTAGAGGGAGGGTCCCAGG 2168
Qy 3059 CCCTGCCAGGAGTCAAGGTGAGGACCAAGCGGGCACCTCACCCAGGACACATTAATTCCA 3118
Db 2169 CTCTGCCAGGAGTTGAGGTGAGGACCAAGCAGGCTCCGCATCCAGGACACATGGGTTCCA 2228
Qy 3119 ATGAATTTTGATATCTCTTGCTGCCCTTC-CCCAAGGACCTAGGCACGTGTGGCCAGATG 3177
Db 2229 ATGAATTTCGACATCTTTTGCTGTCGTTCTTCGGAAGACCTAGGCACAGGTGGCCAGATG 2288
Qy 3178 TTTGTCCCCTCCTGTCCTTCCATTCCTTATCATGGATGTGAACTCTTGATTTGGATTTCT 3237
Db 2289 TGGGGTTTCTTAGGTCCT---GTTCCCTCTCAGGCATGTGAGCTCTTGATCTGAGTTTCT 2345
Qy 3238 CAGACCAGCAAAAGGGCAGGATCCAGGCCCTGCCAGGAAAAATATAAGGGCCCTGCGTGA 3297
Db 2346 CAGGCCAGCAAAAGAGTGGGATCCAGGCCCTGCCTGGAGAAATGTGAGGGCCCTGAGTGA 2405
Qy 3298 GAACAGAGGGGGTCATCCACTGCATGAGAGTGGGGATGTCACAGAGTCCAGCCCACCCTC 3357
Db 2406 ACACAGTGGGGATCATCCACTCCATGAGAGTGGGGACCTCACAGAGTCCAGCCTACCCTC 2465
Qy 3358 CTGGTAGCACTGAGAAGCCAGGGCTGTGCTTGCGGTCTGCACCCTGAGGGCCCGTGGATT 3417
Db 2466 TTGATGGCACTGAGGGACCGGGGCTGTGCTTACAGTCTGCACCCTAAGGGCCCATGGATT 2525
Qy 3418 CCTCTTCCTGGAGCTCCAGGAACCAGGCAGTGAGGCCTTGGTCTGAGACAGTATCCTCAG 3477
Db 2526 CCTCTCCTAGGAGCTCCAGGAACAAGGCAGTGAGGCCTTGGTCTGAGACAGTGTCCTCAG 2585
Qy 3478 GTCACAGAGCAGAGGATGCACAGGGTGTGCCAGCAGTGAATGTTTGCCCTGAATGCACAC 3537
Db 2586 GTTACAGAGCAGAGGATGCACAGGCTGTGCCAGCAGTGAATGTTTGCCCTGAATGCACAC 2645
Qy 3538 CAAGGGCCCCACCTGCCACAGGACACATAGGACTCCACAGAGTCTGGCCTCACCTCCCTA 3597
Db 2646 CAAGGGCCCCACCTGCCACAAGACACATAGGACTCCAAAGAGTCTGGCCTCACCTCCCTA 2705
Qy 3598 CTGTCAGTCCTGTAGAATCGACCTCTGCTGGCCGGCTGTACCCTGA-GTACCCTCTCACT 3656
Db 2706 CCATCAATCCTGCAGAATCGACCTCTGCTGGCCGGCTATACCCTGAGGTGCTCTCTCACT 2765
Qy 3657 TCCTCCTTCAGGTTTTCAGGGGACAGGCCAACCCAGAGGACAGGATTCCCTGGAGGCCAC 3716
Db 2766 TCCTCCTTCAGGTTCTGAGCAGACAGGCCAA-CCGGAGGACAGGATTCCCTGGAGGCCAC 2824
Qy 3717 AGAGGAGCACCAAGGAGAAGATCTGTAAGTAGGCCTTTGTTAGAGTCTCCAAGGTTCAGT 3776
Db 2825 AGAGGAGCACCAAGGAGAAGATCTGTAAGTAAGCCTTTGTTAGAGCCTCTAAGATTTGGT 2884
Qy 3777 TCTCAGCTGAGGCCTCTCACACACTCCCTCTCTCCCCAGGCCTGTGGGTCTTCATTGCCC 3836
Db 2885 TCTCAGCTGAGGTCTCTCACATGCTCCCTCTCTCCGTAGGCCTGTGGGTCCCCATTGCCC 2944
Qy 3837 AGCTCCTGCCCACACTCCTGCCTGCTGCCCTGACGAGAGTCATCATGTCTCTTGAGCAGA 3896
Db 2945 AGCTTTTGCCTGCACTCTTGCCTGCTGCCCTGAGCAGAGTCATCATGTCTTCTGAGCAGA 3004
Qy 3897 GGAGTCTGCACTGCAAGCCTGAGGAAGCCCTTGAGGCCCAACAAGAGGCCCTGGGCCTGG 3956
Db 3005 AGAGTCAGCACTGCAAGCCTGAGGAAGGCGTTGAGGCCCAAGAAGAGGCCCTGGGCCTGG 3064
Qy 3957 TGTGTGTGCAGGCTGCCAC----------CTCCTCCTCCTCTCCTC 3992
Db 3065 TGGGTGCGCAGGCTCCTACTACTGAGGAGCAGGAGGCTGCTGTCTCCTCCTCCTCTCCTC 3124
Qy 3993 TGGTCCTGGGCACCCTGGAGGAGGTGCCCACTGCTGGGTCAACAGATCCTCCCCAGAGTC 4052
Db 3125 TGGTCCCTGGCACCCTGGAGGAAGTGCCTGCTGCTGAGTCAGCAGGTCCTCCCCAGAGTC 3184
Qy 4053 CTCAGGGAGCCTCCGCCTTTCCCACTACCATCAACTTCACTCGACAGAGGCAACCCAGTG 4112
Db 3185 CTCAGGGAGCCTCTGCCTTACCCACTACCATCAGCTTCACTTGCTGGAGGCAACCCAATG 3244
Qy 4113 AGGGTTCCAGCAGCCGTGAAGAGGAGGGGCCAAGCACCTCTTGTATCCTGGAGTCCTTGT 4172
Db 3245 AGGGTTCCAGCAGCCAAGAAGAGGAGGGGCCAAGCACCTCGCCTGACGCAGAGTCCTTGT 3304
Qy 4173 TCCGAGCAGTAATCACTAAGAAGGTGGCTGATTTGGTTGGTTTTCTGCTCCTCAAATATC 4232
Db 3305 TCCGAGAAGCACTCAGTAACAAGGTGGATGAGTTGGCTCATTTTCTGCTCCGCAAGTATC 3364
Qy 4233 GAGCCAGGGAGCCAGTCACAAAGGCAGAAATGCTGGAGAGTGTCATCAAAAATTACAAGC 4292
Db 3365 GAGCCAAGGAGCTGGTCACAAAGGCAGAAATGCTGGAGAGAGTCATCAAAAATTACAAGC 3424
Qy 4293 ACTGTTTTCCTGAGATCTTCGGCAAAGCCTCTGAGTCCTTGCAGCTGGTCTTTGGCATTG 4352
Db 3425 GCTGCTTTCCTGTGATCTTCGGCAAAGCCTCCGAGTCCCTGAAGATGATCTTTGGCATTG 3484
Qy 4353 ACGTGAAGGAAGCAGACCCCACCGGCCACTCCTATGTCCTTGTCACCTGCCTAGGTCTCT 4412
Db 3485 ACGTGAAGGAAGTGGACCCCACCAGCAACACCTACACCCTTGTCACCTGCCTGGGCCTTT 3544
Qy 4413 CCTATGATGGCCTGCTGGGTGATAATCAGATCATGCCCAAGACAGGCTTCCTGATAATTG 4472
Db 3545 CCTATGATGGCCTGCTGGGTAATAATCAGATCTTTCCCAAGACAGGCCTTCTGATAATCG 3604
Qy 4473 TCCTGGTCATGATTGCAATGGAGGGCGGCCATGCTCCTGAGGAGGAAATCTGGGAGGAGC 4532
Db 3605 TCCTGGGCACAATTGCAATGGAGGGCGACAGCGCCTCTGAGGAGGAAATCTGGGAGGAGC 3664
Qy 4533 TGAGTGTGATGGAGGTGTATGATGGGAGGGAGCACAGTGCCTATGGGGAGCCCAGGAAGC 4592
Db 3665 TGGGTGTGATGGGGGTGTATGATGGGAGGGAGCACACTGTCTATGGGGAGCCCAGGAAAC 3724
Qy 4593 TGCTCACCCAAGATTTGGTGCAGGAAAAGTACCTGGAGTA-CGGCAGGTGCCGGACAGTG 4651
Db 3725 TGCTCACCCAAGATTGGGTGCAGGAAAACTACCTGGAGTACCGGCAGGTACCCGGCAGTA 3784
Qy 4652 ATCCCGCACGCTATGAGTTCCTGTGGGGTCCAAGGGCCCTCGCTGAAACCAGCTATGTGA 4711
Db 3785 ATCCTGCGCGCTATGAGTTCCTGTGGGGTCCAAGGGCTCTGGCTGAAACCAGCTATGTGA 3844
Qy 4712 AAGTCCTTGAGTATGTGATCAAGGTCAGTGCAAGAGTTCGCTTTTTCTTCCCATCCCTGC 4771
Db 3845 AAGTCCTGGAGCATGTGGTCAGGGTCAATGCAAGAGTTCGCATTGCCTACCCATCCCTGC 3904
Qy 4772 GTGAAGCAGCTTTGAGAGAGGAGGAAGAGGGAGTCTGAGCATGAGTTGCAGCCAAGGCCA 4831
Db 3905 GTGAAGCAGCTTTGTTAGAGGAGGAAGAGGGAGTCTGAGCATGAGTTGCAGCCAGGGCTG 3964
Qy 4832 GTGGGAGGG----GGACTGGGCCAGTGCACCTTCCAGGGCCGCGTCCAGCAGCTTCCCCT 4887
Db 3965 TGGGGAAGGGGCAGGGCTGGGCCAGTGCATCTAACA--GCCCTGTGCAGCAGCTTCCCTT 4022
Qy 4888 GCCTCGTGTGACATGAGGCCCATTCTTCACTC----TGAAGAGAGCGGTCAGTGTTCTCA 4943
Db 4023 GCCTCGTGTAACATGAGGCCCATTCTTCACTCTGTTTGAAGAAAATAGTCAGTGTTCTTA 4082
Qy 4944 GTAGTAGGTTTCTGTTCTATTGGGTGACTTGGAGATTTATCTTTGTTCTCTTTTGGAATT 5003
Db 4083 GTAGTGGGTTTCTATTTTGTTGGATGACTTGGAGATTTATCTCTGTTTCCTTTTACAATT 4142
Qy 5004 GTTCAAATGTTTTTTTTTAAGGGATGGTTGAATGAACTTCAGCATCCAAGTTTATGAATG 5063
Db 4143 GTTGAAATG-TTCCTTTTAATGGATGGTTGAATTAACTTCAGCATCCAAGTTTATGAATC 4201

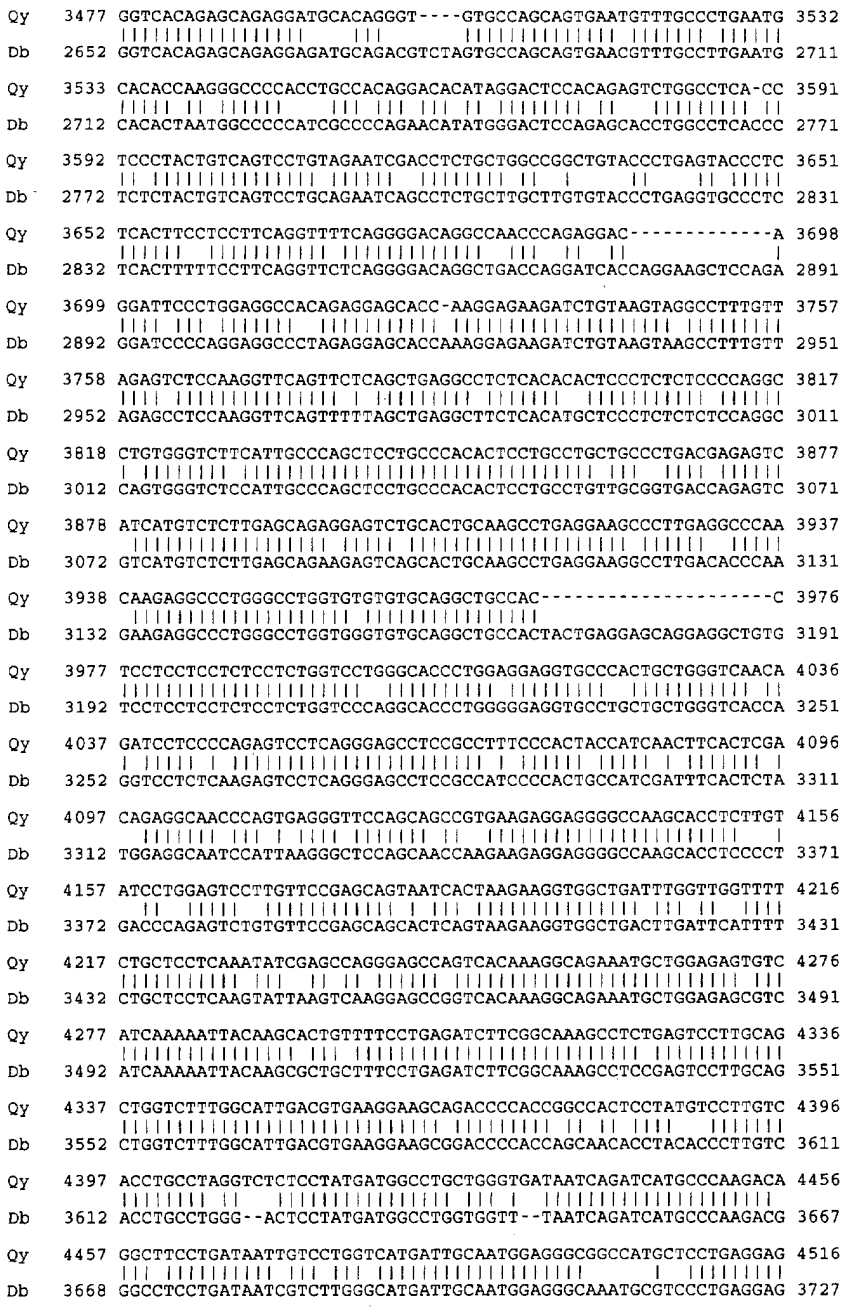
```
Qy   5064 ACAGCAGTCACACAGTTCTGTGTATATAGTTTAAGGGTAAGAGTCTTGTGTTTTATTCAG 5123
Db   4202 GTAGTTAACGTATATTGCTGTTAATATAGTTTAGGAGTAAGAGTCTTGTTTTTTATTCAG 4261

Qy   5124 ATTGGGAAATCCATTCTATTTTGTGAATTGGG---ATAATAACAGCAGTGGAATAAGTAC 5180
Db   4262 ATTGGGAAATCCGTTCTATTTTGTGAATTTGGGACATAATAACAGCAGTGGAGTAAGTAT 4321

Qy   5181 TTAGAAATGTGAAAAATGAGCAGTAAAATAGATGAGATAAAGAACTAAAGAAATTAAGAG 5240
Db   4322 TTAGAAGTGTG---AATTCACCGTGAAATAGGTGAGAT----------AAATTAAAAG 4366

Qy   5241 ATAGTCAATTCTTGCCTTATACCTCAGTCTATTCTGTAAAATTTTTAAAGATATATGCAT 5300
Db   4367 ATACTTAATTCCCGCCTTATGCCTCAGTCTATTCTGTAAAATTTAAAAATATATATGCAT 4426

Qy   5301 ACCTGGATTTCCTTGGCTTCTTTGAGAATGTAAGAGAAATTAAATCTGAATAAAGAATTC 5360
Db   4427 ACCTGGATTTCCTTGGCTTC---GTGAATGTAAGAGAAATTAAATCTGAATAAATAATTC 4483

Qy   5361 TTCCTGTTCACTGGCTCTTTTCTTCTCCATGCACTGAGCATCTGCTTTTTGGAAGGCCCT 5420
Db   4484 TTTCTGTTAACTGGCTCATTTCTTCTCTATGCACTGAGCATCTGCTCTGTGGAAGGCCCA 4543

Qy   5421 GGGTTAGTAGTGGAGATGCTAAGGTAAGCCAGACTCATACCCACCCATAGGGTCGTAGAG 5480
Db   4544 GGATTAGTAGTGGAGATACTAGGGTAAGCCAGACACACACCTACCGATAGGGTATTAAGA 4603

Qy   5481 TCTAGGAGCTGCAGTCACGTAATCGAGGTGGCAAGATGTCCTCTAAAGATGTAGGGAAAA 5540
Db   4604 GTCTAGGAGCGCGGTCATATAATTAAGGTGACAAGATGTCCTCTAAGATGTAGGGGAAAA 4663

Qy   5541 GTGAGAGAGGGGTGAGGGTGTGGGGCTCCGGGTGAGAGTGGTGGAGTGTCAATGCCCTGA 5600
Db   4664 GT----AACGAGTGTGGGTATGGGGCTCCAGGTGAGAGTGGTCGGGTGTAAATTCCCTGT 4719

Qy   5601 GCTGGGGCATTTTGGGCTTTGGGAAACTGCAGTTCCTTCTGGGGGAGCTGATTGTAATGA 5660
Db   4720 G-TGGGGCCTTTTGGGCTTTGGGAAACTCCATTTTCTTCTGAGGGATCTGATTCTAATGA 4778

Qy   5661 TCTTGGGTGGATCC 5674
Db   4779 AGCTTGGTGGGTCC 4792
```

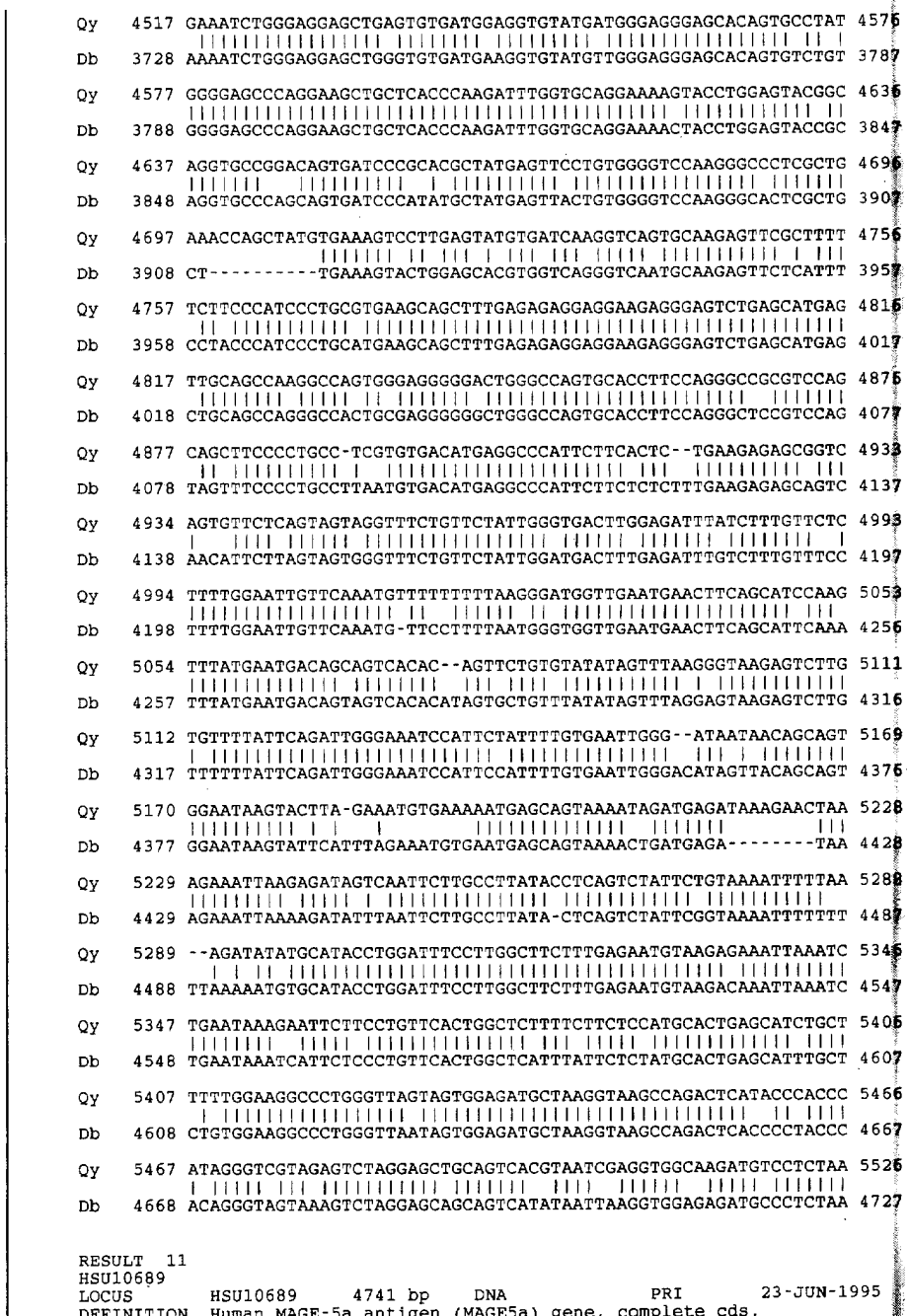
```
RESULT  10
HSU10690
LOCUS       HSU10690     4736 bp    DNA             PRI       23-JUN-1995
DEFINITION  Human MAGE-5b antigen (MAGE5b) gene, complete cds.
ACCESSION   U10690
VERSION     U10690.1  GI:533520
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4736)
  AUTHORS   De Plaen,E., Arden,K., Traversari,C., Gaforio,J.J., Szikora,J.P.,
            De Smet,C., Brasseur,F., van der Bruggen,P., Lethe,B., Lurquin,C.
            Brasseur,R., Chomez,P., De Backer,O., Cavenee,W. and Boon,T.
  TITLE     Structure, chromosomal localization, and expression of 12 genes of
            the MAGE family
  JOURNAL   Immunogenetics 40 (5), 360-369 (1994)
  MEDLINE   95012457
REFERENCE   2  (bases 1 to 4736)
  AUTHORS   De Plaen,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JUN-1994) Etienne De Plaen, Ludwig Institute for
            Cancer Research, 74 Avenue Hippocrate, Brussels, 1200, Belgium
FEATURES             Location/Qualifiers
     source          1. .4736
                     /organism="Homo sapiens"
                     /isolate="patient MZ2"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /sex="female"
                     /cell_type="lymphocyte"
                     /tissue_type="blood"
                     /dev_stage="adult"
     exon            644. .696
                     /number=1
     exon            2848. .2934
                     /number=2
     exon            3010. .>4555
                     /number=3
     gene            3075. .3449
                     /gene="MAGE5b"
     CDS             3075. .3449
                     /gene="MAGE5b"
                     /codon_start=1
                     /product="MAGE-5b antigen"
                     /protein_id="AAA68874.1"
                     /db_xref="GI:533521"
                     /translation="MSLEQKSQHCKPEEGLDTQEEALGLVGVQAATTEEQEAVSSSSP
                     LVPGTLGEVPAAGSPGPLKSPQGASAIPTAIDFTLWRQSIKGSSNQEEEGPSTSPDPE
                     SVFRAALSKKVADLIHFLLLKY"
BASE COUNT     1114 a   1278 c   1304 g   1040 t
ORIGIN
```

```
  Query Match             42.8%;  Score 2429.6;  DB 10;  Length 4736;
  Best Local Similarity   75.9%;  Pred. No. 0;
  Matches 3736;  Conservative    0;  Mismatches  869;  Indels  315;  Gaps   47;

Qy    715 TTCCCGCCAGGAAACATCCGGGTGCCCGGATGTGACGCCACTGACTTGCGCATTGTGGGG 774
Db     15 TTCAACCCAGGGAATCCCTGGGTGACCAGATGTGGTGCCACTGTCTTGCACATTTGAGGT 74

Qy    775 CAGAGAGAAGCGAGGTTTCCATTCTGAGGGACGGCGTAGAGTTCGGCCGAAGGAACCTGA 834
Db     75 CGGAGAGAAGCAAGGGCCTCGCTCTCAGGGGCAGC-TGGAGATCAGCTGAGGGCAGCTGG 133

Qy    835 CCCAGGCTCTGTGAGGAGGCAAGGTGAGAGGCTGAGGGAGGACTGAGGACCCCGCCACTC 894
Db    134 CCCTGGCTCTGTGAGGATGCAAGGTGAGATGCTGAGGGAGGACTAAGGAGTATCCCACCC 193

Qy    895 CAAATAGAGAGCCCCAAATATTCCAG---CCCCGCCCTTGCTGCCAGCCCTGGCCCACCC 951
Db    194 CTGGTAGTGGACCCCAAATAATCCAGTGCCACCTCTCCTGCTGCTAGCTCTGGACCATCC 253

Qy    952 GCGGGAAGACGTCTCAGCCTGGGCTGCCCCCAGACCCCTGCTCCAAAAGCCTTGAGAGAC 1011
Db    254 AGGGCAGGACTTCTTAGGCTGGGCCACCCCCAGTCCCCCACCGCTTAAGCCGCAGGGGA- 312

Qy   1012 ACCAGGTTCTTCTCCCCAAGCTCTGGAATCAGAGGTTGCTGTGACCAGGGCAGGACTGGT 1071
Db    313 -------------------CTCAGGAGACAGAGCTTGGTATGACCAGGGCAGGACTGGT 352

Qy   1072 TAGGAGAGGGCAGGGCACAGGCTCTGCCAGGCATCAAGATCAGCACCCAAGAGGGAGGGC 1131
Db    353 TAGGAGAGGACAGCTCCCA-GCTCTGCCAGGAAACAACGTCAGGAACCTAAGGGAAAGCT 411

Qy   1132 TGTGGGCCCCCAAGACTGCACTCCAATCCCCACTCCCACCCCATTCGCATTCCCATTCCC 1191
Db    412 GAGGCTACCCC------------------------------------------------- 422

Qy   1192 CACCCAACCCCCATCTCCTCAGCTACACCTCCACCCCCATCCCTACTCCTACTCCGTCAC 1251
Db    423 ------------------------CACCCCAAACTCTATTCCTGTCCCTACCTCCGTCC 457

Qy   1252 CTGACCACCACCCTCCAGCCCCAGCACCAGCCCCAACCCTTCTGCCACCTCACCCTCACT 1311
Db    458 CCCACCTACACCCCCCATTCC-----CCCACCCCTTCCCTACCGGCACCTCTATCCCACA 512

Qy   1312 GCCCCCAACCCCACCCTCATCTCTCTCATGTGCCCCACTCCCATCGCCTCCCCCATTCTG 1371
```


Db 513 TCCCCCA----------CCCCTATCCTG 530
Qy 1372 GCAGAATCCGGTT-TGCCCCTGCTCTCAACCCAGGGAAGCCCTGGTAGGCCCGATGTGAA 1430
Db 531 GCAGAATCCGATTCTGCCCCTGATTTCAACCCAGGGAAGCCCTAGGGGGCCGGATGTGAT 590
Qy 1431 ACCACTGACTTGAACCTCACAGATCTGAGAGAAGCCAGGTTCATTTAATGGTTCTGAGGG 1490
Db 591 GCTGCTGACTTGTGCATTGGGGGTCAGAGAGAATCAAGG------GCATGGTTCTGAGAA 644
Qy 1491 GCGGCTTGAGATCCACTGAGGGGAGTGGTTTTAGGCTCTGTGAGGAGGCAAGGTGAGATG 1550
Db 645 GCCGACTGAGATCAGCAGAGGGGAATGGGCCCGGGCTCTGTGAGGAGGCAAGGTGAGACC 704
Qy 1551 CT-GAGGGAGGACTGAGGAGGCACACACCCCAGGTAGATGGCCCCAAAATGATCCAGTAC 1609
Db 705 CCCGAGGAAGGAATGAGGAAGCCCTCACCC--AGATAGAGAACCCCAAATAATCCAGTAC 762
Qy 1610 CACCCCTGCTGCCAGCCCTGGACCACCCGGCCAGGACAGATGTCTCAGCTGGACCACCCC 1669
Db 763 TACCTCTGCTGCCAGCCCTGGACCAC----CCAGGGCAGACTTCTCAGGCTGAACCTTCC 818
Qy 1670 CCGTCCCGTCCCACTGCCACTTAACCCACAGGGCAATCTGTAGTCATAGCTT-ATGTGAC 1728
Db 819 C---CCCTCCCCACTGCCACTTAAGCCACAAGGGACTCTGGAGTCAGACCTTGGTGTGAC 875
Qy 1729 CGGGGCAGGGTTGGTCAGGAGAGGCAGGGCCCAGGCATCAAGGTCCAGCATCCGCCCGGC 1788
Db 876 CAGGGAAGGGCCGGTCAGGAGAGG------------GCAGGGGCCAGGCTCTGTCAGGC 922
Qy 1789 ATTAGGGTCAGGACCCTGGGAGGGAACTGAGGGTTCCCCACCCACACCTGTCTCCTCATC 1848
Db 923 ATCAAAATCAGGACCCTGAGAGAGAATTGAGGGCCCCCACCCCAACCCCTATACCCATCC 982
Qy 1849 TCCACCGCCACCCCACTCACATTCCCATACCTACC----CCCTACCCCCAACCTCATCTT 1904
Db 983 CTAACCCCATACCCACTCTACTTGCATTCCCAGCCCCATCCCCACACCCTACCCCATCTT 1042
Qy 1905 GTCAGAA----------TCCCTGCTGTCAACCCACGGAAGCCACGGGAATGGCGGCCAGG 1954
Db 1043 GGCAGAATCTGTTTCTTTCCCTGCAGTCAACCCACAGAAGCCCCAGGAATGACAGACAGG 1102
Qy 1955 CACTCGGATCTTGACGTCCCCATCCAGGGTCTGATGGAGGGAAGGG-------GCTTGAA 2007
Db 1103 CACACCCATTCTGACGTCCACATCCAGGGCTGAAGGAGGGAAAGGGCTTAGTATCATGAG 1162
Qy 2008 CAGGGCCTCAGGGGAGCAGAGG----GAGGGCCCTACTGCGAGATGAGGGAGGCCTCAGA 2063
Db 1163 CAGGGCCTCAGGGGAGTCTCTGCTCCTCAAGCCCTGCTGGGAGTAAAGGGAGGCCTCAGG 1222
Qy 2064 GGACCCAGCACCCTAGGA---------CACCGCACCCCTGTCTGAGACTGAGGC--TGCC 2112
Db 1223 GAACCCAGGTCCTCAGGATAGGGGGTCCACTCCAACCCTGTCTGAGACTGAGGCGCCTCC 1282
Qy 2113 ACTTCTGGCCTCAAGAATCAGAACGATGGGGACTCAGATTGCATGGGGGTGGGACCCAGG 2172
Db 1283 TCTTTCATCCTCGGGAATCACAGGGATGGAGACTCACGTCAGCAGAGGGTGGGGCCCAAC 1342
Qy 2173 CCTGCAAGGCTTACGCGGAGGAAGAGGAGGGAGGACTCAGGGGACCTTGGAATCCAGATC 2232
Db 1343 CCTGCCAGGATCAAGGAGAGGAAGAAGAGGGAGGACTCAGGGTACCTTTGAGTCCAGAAC 1402
Qy 2233 AGTGTGGACCTCGGCCCTGAGAGGTCCAGGGCACGGTGGCCACATATGGCCCATATTTCC 2292
Db 1403 AATGGGGACCTTTGCCCTGGGAGGTCCAGTGCACAGTGGCCACCTGTAGCCCATGCTTGC 1462
Qy 2293 TGCATCTTTGAGGTGAC----AGGACAGAGCTGTGGTCTGAGAAGTGGGGCCTCAGGTCA 2348
Db 1463 TGCACCTTCTGGGTGACAAAGAGGAGAGGGCTGTGGTCAGAGCAGTGGTGACTCAGGTCA 1522
Qy 2349 ACAGAGGGAGGAGTTCCAGGATCCATATGGCCCAAGATGTGCCCCCTTCATGAGGACTGG 2408


Db 1523 GCAGAGGGAGGAGTCCCAGCATCTGCAGGCCCCAATGTGTGCCCCATTCATGAAGATTGG 1582
Qy 2409 GGATATCCCCGGCTCAGAAAGAAGGGACTCCACACAGTCTGGCTGTCCCCTTTTAGTAGC 2468
Db 1583 GGACA-CCTTGGCTCAGAAAGAAGGGACCCCACAGAGTCTGGCTGTCCCCTGATTTTTGC 1641
Qy 2469 TCTAGGGGGACCAGATCAGGGATGGCGGTATGTTCCATTCTCACTTGTACCACAGGCAGG 2528
Db 1642 TCAGAGGGGACCAAATCAAGGATAGCCCTATGTGCCAACCTCATTTGTGCCACAGGAAAG 1701
Qy 2529 AAGTTGGGGGGCCCTCAGGGAGATGGGGTCTTGGGGTAAAGGGGGGATGTCTACTCATGT 2588
Db 1702 AAGTTGAAGAGCCCTCAGGGTGATGGGGTCTTGCAGTAAAGGGGAGCTATCTGCTCATCT 1761
Qy 2589 CAGGGAATTGGGGGTTGAGGAAGCACAGGCGCTGGCAGGAATAAAGATGAGTGAGACAGA 2648
Db 1762 CAGGGGGTTTCAGGTTGAGGAATGGCAGGCCCCATCACGATGAAGAGTAACCCACAGG-- 1819
Qy 2649 CAAGGCTATTGGAATCCACACCCCAGAACCAAAGGGGTCAGCCCTGGACACCTCA----- 2703
Db 1820 ---AGCCATAGAAACACTCACCCCAGAACCAAAGGGGTCATACCTGGACACCCCATGTGG 1875
Qy 2704 ----CCCAGGATGTGGCTTCTTTTTCACTCCTGTTTCCAGATCTGGGGCAGGTGAGGACC 2759
Db 1877 GGGTGACAGGATGTAGCT-CCATCTCATTCCTGTTTTCAGATCTCGGGGAGGTGAGGAAC 1935
Qy 2760 TCATTCTCAGAGGGTGACTCAGGTCAACGTAGGGACCCCCATCTGGTCTAAAGACAGAGC 2819
Db 1936 TTGTTCTCCGAGGATGACTCAGGTCAACACAGGGGCCCCCATCTGGTGGATAGACAGAGT 1995
Qy 2820 GGTCCCAGGATCTGCCATGCGTTCGGGTGAGGAACATGAGGGAGGACTGAGGGTACCCCA 2879
Db 1996 GGTCCCAGGATCTGTCAGTAGTTCCGGTGAGGAACATGAGGGACGATTGAGGGCACCCTT 2055
Qy 2880 GGACCAGAACACTGAGGGAGACTGCACAGAAATCAGCCCTGCCCCTGCTGTCACCCCAGA 2939
Db 2056 GGGCCAGAACACAGATGAGGACCTCACGGAAATCTGCCCTGCCCCTGCTGTCACTCCAGA 2115
Qy 2940 GAGCATGGGCTGGGCCGTCTGCCGAGGTCCTTCC--GTTATCCTGGGATCATTGATGTCA 2997
Db 2116 GAGCATGGGCAGGGCTGTCTGCTGCAGTCCCCCCCACTTACCCTGGGATCATTGGTGTCA 2175
Qy 2998 GGGACGGGGAGGCCTTGGTCTGAGAAGGCTGCGCTCAGGTCAGTAGAGGGAGCGTCCCAG 3057
Db 2176 GTGATGGGGAGGTCTTTGTC-GAGGGGTCTGCACTCAGGTCAGTAGAGGGAGCGTCTTAG 2234
Qy 3058 GCCCTGCCAGGAGTCAAGGTGAGGACCAAGCGGGCACCTCACCCAGGACACATTAATTCC 3117
Db 2235 GCCCTGCCAGGAGACAAGGTAAGAACGAAGCAGGTTCCTCACCCAGGACACATGAATTCC 2294
Qy 3118 AATGAATTTTGATATCTCTTGCTGCCCTTCCCCA-AGGACCTAGGCACGTGTGGCCAGAT 3176
Db 2295 AATGCATTTCAGCATCTCTTCCTGTCCTTCCCAAGAGGACCTGGGCACGTGTGGCCAGAT 2354
Qy 3177 GTTTGTCCCCTCCTGTCCTTCCATTCCTTATCATGGATGTGAACTCTTGATTTGGATTTC 3236
Db 2355 GTGAGTCTCCTCATGTCCT---GTTCCCTATCAGGGATGTGAGCTCTTAATCTGAGTTTC 2411
Qy 3237 TCAGACCAGCAAAAGGGCAGGATCCAGGCCCTGCCAGGAAAAATATAAGGGCCCTGCGTG 3296
Db 2412 TCAGGCCAGCAAAAGGGTGGGATCCAGGCCTTGCCAGGAGAAAGGTGAGGGCCCTGTGTG 2471
Qy 3297 AGAACAGAGGGGGTCATCCACTGCATGAGAGTGGGGATGTCACAGAGTCCAGCCCACCCT 3356
Db 2472 AGCACAGAGGGGACCATTCACCCCAAGAGGGTGGAGACCTCACAGATTCCAGCCTACCCT 2531
Qy 3357 CCTGGTAGCACTGAGAAGCCAGGGCTGTGCTTGCGGTCTGCACCCTGAGGGCCCGTGGAT 3416
Db 2532 CCTGTTAGCACTGGGGGCCTGAGGCTGTGCTTGCAGTCTGCACCCTGAGGGCCCATGCAT 2591
Qy 3417 TCCTCTTCCTGGAGCTCCAGGAACCAGGCAGTGAGGCCTTGGTCTGAGACAGTATCCTCA 3476
Db 2592 TCCTCTTCCAGGAGCTCCAGGAAACAGACACTGAGGCCTTGGTCTGAGGCCGTGCCCTCA 2651

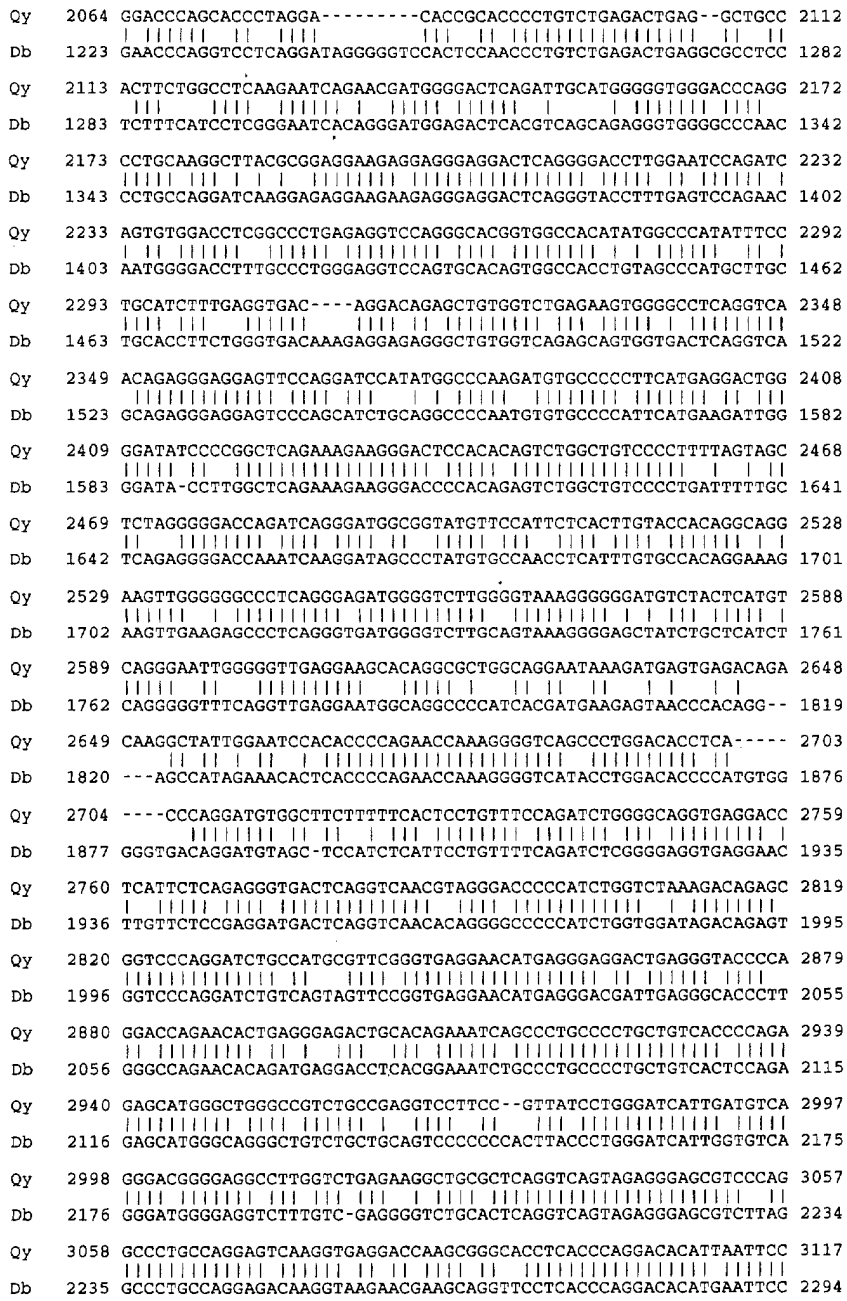




RESULT 11
HSU10689
LOCUS HSU10689 4741 bp DNA PRI 23-JUN-1995
DEFINITION Human MAGE-5a antigen (MAGE5a) gene, complete cds.

```
ACCESSION   U10689
VERSION     U10689.1  GI:533518
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4741)
  AUTHORS   De Plaen,E., Arden,K., Traversari,C., Gaforio,J.J., Szikora,J.P.,
            De Smet,C., Brasseur,F., van der Bruggen,P., Lethe,B., Lurquin,C.
            Brasseur,R., Chomez,P., De Backer,O., Cavenee,W. and Boon,T.
  TITLE     Structure, chromosomal localization, and expression of 12 genes of
            the MAGE family
  JOURNAL   Immunogenetics 40 (5), 360-369 (1994)
  MEDLINE   95012457
REFERENCE   2  (bases 1 to 4741)
  AUTHORS   De Plaen,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JUN-1994) Etienne De Plaen, Ludwig Institute for
            Cancer Research, 74 Avenue Hippocrate, Brussels, 1200, Belgium
FEATURES             Location/Qualifiers
     source          1. .4741
                     /organism="Homo sapiens"
                     /isolate="patient MZ2"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /sex="female"
                     /cell_type="lymphocyte"
                     /tissue_type="blood"
                     /dev_stage="adult"
     exon            644. .696
                     /number=1
     exon            2848. .2934
                     /number=2
     exon            3010. .>4560
                     /number=3
     gene            3075. .3449
                     /gene="MAGE5a"
     CDS             3075. .3449
                     /gene="MAGE5a"
                     /codon_start=1
                     /product="MAGE-5a antigen"
                     /protein_id="AAA68873.1"
                     /db_xref="GI:533519"
                     /translation="MSLEQKSQHCKPEEGLDTQEEALGLVGVQAATTEEQEAVSSSSP
                     LVPGTLGEVPAAGSPGPLKSPQGASAIPTAIDFTLWRQSIKGSSNQEEEGPSTSPDPE
                     SVFRAALSKKVADLIHFLLLKY"
BASE COUNT     1118 a   1274 c   1305 g   1044 t
ORIGIN
```

```
  Query Match           42.7%;  Score 2422.8;  DB 10;  Length 4741;
  Best Local Similarity  75.7%;  Pred. No. 0;  .
  Matches 3726;  Conservative    0;  Mismatches  882;  Indels  314;  Gaps   46;

Qy    715 TTCCCGCCAGGAAACATCCGGGTGCCCGGATGTGACGCCACTGACTTGCGCATTGTGGGG 774
Db     15 TTCAACCCAGGGAATCCCTGGGTGACCAGATGTGGTGCCACTGTCTTGCACATTTGAGGT 74

Qy    775 CAGAGAGAAGCGAGGTTTCCATTCTGAGGGACGGCGTAGAGTTCGGCCGAAGGAACCTGA 834
Db     75 CGGAGAGAAGCAAGGGCCTCGCTCTCAGGGGCAGC-TGGAGATCAGCTGAGGGCAGCTGG 133

Qy    835 CCCAGGCTCTGTGAGGAGGCAAGGTGAGAGGCTGAGGGAGGACTGAGGACCCCGCCACTC 894
Db    134 CCCTGGCTCTGTGAGGATGCAAGGTGAGATGCTGAGGGAGGACTAAGGAGTATCCCACCC 193

Qy    895 CAAATAGAGAGCCCCAAATATTCCAG---CCCCGCCCTTGCTGCCAGCCCTGGCCCACCC 951
Db    194 CTGGTAGTGGACCCCAAATAATCCAGTGCCACCTCTCCTGCTGCTAGCTCTGGACCATCC 253

Qy    952 GCGGGAAGACGTCTCAGCCTGGGCTGCCCCCAGACCCCTGCTCCAAAAGCCTTGAGAGAC 1011
Db    254 AGGGCAGGACTTCTTAGGCTGGGCCACCCCCAGTCCCCCACCGCTTAAGCCGCAGGGGA- 312

Qy   1012 ACCAGGTTCTTCTCCCCAAGCTCTGGAATCAGAGGTTGCTGTGACCAGGGCAGGACTGGT 1071
Db    313 -------------------CTCAGGAGACAGAGCTTGGTATGACCAGGGCAGGACTGGT 352

Qy   1072 TAGGAGAGGGCAGGGCACAGGCTCTGCCAGGCATCAAGATCAGCACCCAAGAGGGAGGGC 1131
Db    353 TAGGAGAGGACAGCTCCCA-GCTCTGCCAGGAAACAACGTCAGGAACCTAAGGGAAAGCT 411

Qy   1132 TGTGGGCCCCCAAGACTGCACTCCAATCCCCACTCCCACCCCATTCGCATTCCCATTCCC 1191
Db    412 GAGGCTACCCC---------------------------------------------------- 422

Qy   1192 CACCCAACCCCCATCTCCTCAGCTACACCTCCACCCCCATCCCTACTCCTACTCCGTCAC 1251
Db    423 ------------------------CACCCCAAACTCTATTCCTGTCCCTACCTCCGTCC 457

Qy   1252 CTGACCACCACCCTCCAGCCCCAGCACCAGCCCCAACCCTTCTGCCACCTCACCCTCACT 1311
Db    458 CCCACCTACACCCCCCATTCC-----CCCACCCCTTCCCTACCGGCACCTCTATCCCACA 512

Qy   1312 GCCCCCAACCCCACCCTCATCTCTCTCATGTGCCCCACTCCCATCGCCTCCCCCATTCTG 1371
Db    513 TCCCCCA----------------------------------------CCCCTATCCTG 530

Qy   1372 GCAGAATCCGGTT-TGCCCCTGCTCTCAACCCAGGGAAGCCCTGGTAGGCCCGATGTGAA 1430
Db    531 GCAGAATCCGATTCTGCCCCTGATTTCAACCCAGGGAAGCCCTAGGGGGCCGGATGTGAT 590

Qy   1431 ACCACTGACTTGAACCTCACAGATCTGAGAGAAGCCAGGTTCATTTAATGGTTCTGAGGG 1490
Db    591 GCTGCTGACTTGTGCATTGGGGGTCAGAGAGAATCAAGG------GCATGGTTCTGAGAA 644

Qy   1491 GCGGCTTGAGATCCACTGAGGGGAGTGGTTTTAGGCTCTGTGAGGAGGCAAGGTGAGATG 1550
Db    645 GCCGACTGAGATCAGCAGAGGGGAATGGGCCCGGGCTCTGTGAGGAGGCAAGGTGAGACC 704

Qy   1551 CT-GAGGGAGGACTGAGGAGGCACACACCCCAGGTAGATGGCCCCAAAATGATCCAGTAC 1609
Db    705 CCCGAGGAAGGAATGAGGAAGCCCTCACCC--AGATAGAGAACCCCAAATAATCCAGTAC 762

Qy   1610 CACCCCTGCTGCCAGCCCTGGACCACCCGGCCAGGACAGATGTCTCAGCTGGACCACCCC 1669
Db    763 TACCTTTGCTGCCAGCCCTGGACCAC----CCAGGGCAGACTTCTCAGGCTGAACCTTCC 818

Qy   1670 CCGTCCCGTCCCACTGCCACTTAACCCACAGGGCAATCTGTAGTCATAGCTT-ATGTGAC 1728
Db    819 C---CCCTCCCCACTGCCACTTAAGCCACAAGGGACTCTGGAGTCAGACCTTGGTGTGAC 875

Qy   1729 CGGGGCAGGGTTGGTCAGGAGAGGCAGGGCCCAGGCATCAAGGTCCAGCATCCGCCCGGC 1788
Db    876 CAGGGAAGGGCCGGTCAGGAGAGG------------GCAGGGGCCAGGCTCTGTCAGGC 922

Qy   1789 ATTAGGGTCAGGACCCTGGGAGGGAACTGAGGGTTCCCCACCCACACCTGTCTCCTCATC 1848
Db    923 ATCAAAATCAGGACCCTGAGAGAGAATTGAGGGCCCCCACCCCAACCCCTATACCCATCC 982

Qy   1849 TCCACCGCCACCCCACTCACATTCCCATACCTACC----CCCTACCCCCAACCTCATCTT 1904
Db    983 CTAACCCCATACCCACTCTACTTGCATTCCCAGCCCCATCCCCACACCCTACCCCATCTT 1042

Qy   1905 GTCAGAA----------TCCCTGCTGTCAACCCACGGAAGCCACGGGAATGGCGGCCAGG 1954
Db   1043 GGCAGAATCTGTTTCTTTCCCTGCAGTCAACCCACAGAAGCCCCAGGAATGACAGACAGG 1102

Qy   1955 CACTCGGATCTTGACGTCCCCATCCAGGGTCTGATGGAGGGAAGGG-------GCTTGAA 2007
Db   1103 CACACCTATTCTGACGTCCACATCCAGGGCTGAAGGAGGGAAAGGGCTTAGTATCATGAG 1162

Qy   2008 CAGGGCCTCAGGGGAGCAGAGG----GAGGGCCCTACTGCGAGATGAGGGAGGCCTCAGA 2063
Db   1163 CAGGGCCTCAGGGGAGTCTCTGCTCCTCAAGCCCTGCTGGGAGTAAAGGGAGGCCTCAGG 1222
```


Qy 2064 GGACCCAGCACCCTAGGA---------CACCGCACCCCTGTCTGAGACTGAG--GCTGCC 2112
Db 1223 GAACCCAGGTCCTCAGGATAGGGGGTCCACTCCAACCCTGTCTGAGACTGAGGCGCCTCC 1282
Qy 2113 ACTTCTGGCCTCAAGAATCAGAACGATGGGGACTCAGATTGCATGGGGGTGGGACCCAGG 2172
Db 1283 TCTTTCATCCTCGGGAATCACAGGGATGGAGACTCACGTCAGCAGAGGGTGGGGCCCAAC 1342
Qy 2173 CCTGCAAGGCTTACGCGGAGGAAGAGGAGGGAGGACTCAGGGGACCTTGGAATCCAGATC 2232
Db 1343 CCTGCCAGGATCAAGGAGAGGAAGAAGAGGGAGGACTCAGGGTACCTTTGAGTCCAGAAC 1402
Qy 2233 AGTGTGGACCTCGGCCCTGAGAGGTCCAGGGCACGGTGGCCACATATGGCCCATATTTCC 2292
Db 1403 AATGGGGACCTTTGCCCTGGGAGGTCCAGTGCACAGTGGCCACCTGTAGCCCATGCTTGC 1462
Qy 2293 TGCATCTTTGAGGTGAC----AGGACAGAGCTGTGGTCTGAGAAGTGGGGCCTCAGGTCA 2348
Db 1463 TGCACCTTCTGGGTGACAAAGAGGAGAGGGCTGTGGTCAGAGCAGTGGTGACTCAGGTCA 1522
Qy 2349 ACAGAGGGAGGAGTTCCAGGATCCATATGGCCCAAGATGTGCCCCCTTCATGAGGACTGG 2408
Db 1523 GCAGAGGGAGGAGTCCCAGCATCTGCAGGCCCCAATGTGTGCCCCATTCATGAAGATTGG 1582
Qy 2409 GGATATCCCCGGCTCAGAAAGAAGGGACTCCACACAGTCTGGCTGTCCCCTTTTAGTAGC 2468
Db 1583 GGATA-CCTTGGCTCAGAAAGAAGGGACCCCACAGAGTCTGGCTGTCCCCTGATTTTTGC 1641
Qy 2469 TCTAGGGGGACCAGATCAGGGATGGCGGTATGTTCCATTCTCACTTGTACCACAGGCAGG 2528
Db 1642 TCAGAGGGGACCAAATCAAGGATAGCCCTATGTGCCAACCTCATTTGTGCCACAGGAAAG 1701
Qy 2529 AAGTTGGGGGGCCCTCAGGGAGATGGGGTCTTGGGGTAAAGGGGGGATGTCTACTCATGT 2588
Db 1702 AAGTTGAAGAGCCCTCAGGGTGATGGGGTCTTGCAGTAAAGGGGAGCTATCTGCTCATCT 1761
Qy 2589 CAGGGAATTGGGGGTTGAGGAAGCACAGGCGCTGGCAGGAATAAAGATGAGTGAGACAGA 2648
Db 1762 CAGGGGGTTTCAGGTTGAGGAATGGCAGGCCCCATCACGATGAAGAGTAACCCACAGG-- 1819
Qy 2649 CAAGGCTATTGGAATCCACACCCCAGAACCAAAGGGGTCAGCCCTGGACACCTCA----- 2703
Db 1820 ---AGCCATAGAAACACTCACCCCAGAACCAAAGGGGTCATACCTGGACACCCCATGTGG 1876
Qy 2704 ----CCCAGGATGTGGCTTCTTTTTCACTCCTGTTTCCAGATCTGGGGCAGGTGAGGACC 2759
Db 1877 GGGTGACAGGATGTAGC-TCCATCTCATTCCTGTTTTCAGATCTCGGGGAGGTGAGGAAC 1935
Qy 2760 TCATTCTCAGAGGGTGACTCAGGTCAACGTAGGGACCCCCATCTGGTCTAAAGACAGAGC 2819
Db 1936 TTGTTCTCCGAGGATGACTCAGGTCAACACAGGGGCCCCCATCTGGTGGATAGACAGAGT 1995
Qy 2820 GGTCCCAGGATCTGCCATGCGTTCGGGTGAGGAACATGAGGGAGGACTGAGGGTACCCCA 2879
Db 1996 GGTCCCAGGATCTGTCAGTAGTTCCGGTGAGGAACATGAGGGACGATTGAGGGCACCCTT 2055
Qy 2880 GGACCAGAACACTGAGGGAGACTGCACAGAAATCAGCCCTGCCCCTGCTGTCACCCCAGA 2939
Db 2056 GGGCCAGAACACAGATGAGGACCTCACGGAAATCTGCCCTGCCCCTGCTGTCACTCCAGA 2115
Qy 2940 GAGCATGGGCTGGGCCGTCTGCCGAGGTCCTTCC--GTTATCCTGGGATCATTGATGTCA 2997
Db 2116 GAGCATGGGCAGGGCTGTCTGCTGCAGTCCCCCCCACTTACCCTGGGATCATTGGTGTCA 2175
Qy 2998 GGGACGGGGAGGCCTTGGTCTGAGAAGGCTGCGCTCAGGTCAGTAGAGGGAGCGTCCCAG 3057
Db 2176 GGGATGGGGAGGTCTTTGTC-GAGGGGTCTGCACTCAGGTCAGTAGAGGGAGCGTCTTAG 2234
Qy 3058 GCCCTGCCAGGAGTCAAGGTGAGGACCAAGCGGGCACCTCACCCAGGACACATTAATTCC 3117
Db 2235 GCCCTGCCAGGAGACAAGGTAAGAACGAAGCAGGTTCCTCACCCAGGACACATGAATTCC 2294
Qy 3118 AATGAATTTTGATATCTCTTGCTGCCCTTCCCCA-AGGACCTAGGCACGTGTGGCCAGAT 3176
Db 2295 AATGCATTTCAGCATCTCTTCCTGTCCTTCCCAAGAGGACCTGGGCACGTGTGGCCAGAT 2354
Qy 3177 GTTTGTCCCCTCCTGTCCTTCCATTCCTTATCATGGATGTGAACTCTTGATTTGGATTTC 3236
Db 2355 GTGAGTCTCCTCATGTCCT---GTTCCCTATCAGGGATGTGAGCTCTTAATCTGAGTTTC 2411
Qy 3237 TCAGACCAGCAAAAGGGCAGGATCCAGGCCCTGCCAGGAAAAATATAAGGGCCCTGCGTG 3296
Db 2412 TCAGGCCAGCAAAAGGGTGGGATCCAGGCCTTGCCAGGAGAAAGGTGAGGGCCCTGTGTG 2471
Qy 3297 AGAACAGAGGGGGTCATCCACTGCATGAGAGTGGGGATGTCACAGAGTCCAGCCCACCCT 3356
Db 2472 AGCACAGAGGGGACCATTCACCCCAAGAGGGTGGAGACCTCACAGATTCCAGCCTACCCT 2531
Qy 3357 CCTGGTAGCACTGAGAAGCCAGGGCTGTGCTTGCGGTCTGCACCCTGAGGGCCCGTGGAT 3416
Db 2532 CCTGTTAGCACTGGGGGCCTGAGGCTGTGCTTGCAGTCTGCACCCTGAGGGCCCATGCAT 2591
Qy 3417 TCCTCTTCCTGGAGCTCCAGGAACCAGGCAGTGAGGCCTTGGTCTGAGACAGTATCCTCA 3476
Db 2592 TCCTCTTCCAGGAGCTCCAGGAAACAGACACTGAGGCCTTGGTCTGAGGCCGTGCCCTCA 2651
Qy 3477 GGTCACAGAGCAGAGGATGCACAGGGT----GTGCCAGCAGTGAATGTTTGCCCTGAATG 3532
Db 2652 GGTCACAGAGCAGAGGAGATGCAGACGTCTAGTGCCAGCAGTGAACGTTTGCCTTGAATG 2711
Qy 3533 CACACCAAGGGCCCCACCTGCCACAGGACACATAGGACTCCACAGAGTCTGGCCTCA-CC 3591
Db 2712 CACACTAATGGCCCCCATCGCCCCAGAACATATGGGACTCCAGAGCACCTGGCCTCACCC 2771
Qy 3592 TCCCTACTGTCAGTCCTGTAGAATCGACCTCTGCTGGCCGGCTGTACCCTGAGTACCCTC 3651
Db 2772 TCTCTACTGTCAGTCCTGCAGAATCAGCCTCTGCTTGCTTGTGTACCCTGAGGTGCCCTC 2831
Qy 3652 TCACTTCCTCCTTCAGGTTTTCAGGGGACAGGCCAACCCAGAGGAC----------A 3698
Db 2832 TCACTTTTTCCTTCAGGTTCTCAGGGGACAGGCTGACCAGGATCACCAGGAAGCTCCAGA 2891
Qy 3699 GGATTCCCTGGAGGCCACAGAGGAGCACC-AAGGAGAAGATCTGTAAGTAGGCCTTTGTT 3757
Db 2892 GGATCCCCAGGAGGCCCTAGAGGAGCACCAAAGGAGAAGATCTGTAAGTAAGCCTTTGTT 2951
Qy 3758 AGAGTCTCCAAGGTTCAGTTCTCAGCTGAGGCCTCTCACACACTCCCTCTCTCCCCAGGC 3817
Db 2952 AGAGCCTCCAAGGTTCAGTTTTTAGCTGAGGCTTCTCACATGCTCCCTCTCTCTCCAGGC 3011
Qy 3818 CTGTGGGTCTTCATTGCCCAGCTCCTGCCCACACTCCTGCCTGCTGCCCCTGACGAGAGTC 3877
Db 3012 CAGTGGGTCTCCATTGCCCAGCTCCTGCCCACACTCCTGCCTGTTGCGGTGACCAGAGTC 3071
Qy 3878 ATCATGTCTCTTGAGCAGAGGAGTCTGCACTGCAAGCCTGAGGAAGCCCTTGAGGCCCAA 3937
Db 3072 GTCATGTCTCTTGAGCAGAAGAGTCAGCACTGCAAGCCTGAGGAAGGCCTTGACACCCAA 3131
Qy 3938 CAAGAGGCCCTGGGCCTGGTGTGTGTGCAGGCTGCCAC----------C 3976
Db 3132 GAAGAGGCCCTGGGCCTGGTGGGTGTGCAGGCTGCCACTACTGAGGAGCAGGAGGCTGTG 3191
Qy 3977 TCCTCCTCCTCTCCTCTGGTCCTGGGCACCCTGGAGGAGGTGCCCACTGCTGGGTCAACA 4036
Db 3192 TCCTCCTCCTCTCCTCTGGTCCCAGGCACCCTGGGGGAGGTGCCTGCTGCTGGGTCACCA 3251
Qy 4037 GATCCTCCCCAGAGTCCTCAGGGAGCCTCCGCCTTTCCCACTACCATCAACTTCACTCGA 4096
Db 3252 GGTCCTCTCAAGAGTCCTCAGGGAGCCTCCGCCATCCCCACTGCCATCGATTTCACTCTA 3311
Qy 4097 CAGAGGCAACCCAGTGAGGGTTCCAGCAGCCGTGAAGAGGAGGGGCCAAGCACCTCTTGT 4156
Db 3312 TGGAGGCAATCCATTAAGGGCTCCAGCAACCAAGAAGAGGAGGGGCCAAGCACCTCCCCT 3371
Qy 4157 ATCCTGGAGTCCTTGTTCCGAGCAGTAATCACTAAGAAGGTGGCTGATTTGGTTGGTTTT 4216

```
Db   3372 GACCCAGAGTCTGTGTTCCGAGCAGCACTCAGTAAGAAGGTGGCTGACTTGATTCATTTT 3431

Qy   4217 CTGCTCCTCAAATATCGAGCCAGGGAGCCAGTCACAAAGGCAGAAATGCTGGAGAGTGTC 4276

Db   3432 CTGCTCCTCAAGTATTAAGTCAAGGAGCTGGTCACAAAGGCAGAAATGCTGGAGAGCGTC 3491

Qy   4277 ATCAAAAATTACAAGCACTGTTTTCCTGAGATCTTCGGCAAAGCCTCTGAGTCCTTGCAG 4336

Db   3492 ATCAAAAATTACAAGCGCTGCTTTCCTGAGATCTTCGGCAAAGCCTCCGAGTCCTTGCAG 3551

Qy   4337 CTGGTCTTTGGCATTGACGTGAAGGAAGCAGACCCCACCGGCCACTCCTATGTCCTTGTC 4396

Db   3552 CTGGTCTTTGGCATTGACGTGAAGGAAGCGGACCCCACCAGCAACACCTACACCCTTGTC 3611

Qy   4397 ACCTGCCTAGGTCTCTCCTATGATGGCCTGCTGGGTG---ATAATCAGATCATGCCCAAG 4453

Db   3612 ACCTGCCTGGG--ACTCCTATGATGGCCTGCTGGTTGATAATAATCAGATCATGCCCAAG 3669

Qy   4454 ACAGGCTTCCTGATAATTGTCCTGGTCATGATTGCAATGGAGGGCGGCCATGCTCCTGAG 4513

Db   3670 ACGGGCCTCCTGATAATCGTCTTGGGCATGATTGCAATGGAGGGCAAATGCGTCCCTGAG 3729

Qy   4514 GAGGAAATCTGGGAGGAGCTGAGTGTGATGGAGGTGTATGATGGGAGGGAGCACAGTGCC 4573

Db   3730 GAGAAAATCTGGGAGGAGCTGAGTGTGATGAAGGTGTATGTTGGGAGGGAGCACAGTGTC 3789

Qy   4574 TATGGGGAGCCCAGGAAGCTGCTCACCCAAGATTTGGTGCAGGAAAAGTACCTGGAGTA- 4632

Db   3790 TGTGGGGAGCCCAGGAAGCTGCTCACCCAAGATTTGGTGCAGGAAAACTACCTGGAGTAC 3849

Qy   4633 CGGCAGGTGCCGGACAGTGATCCCGCACGCTATGAGTTCCTGTGGGGTCCAAGGGCCCTC 4692

Db   3850 CGGCAGGTGCCCAGCAGTGATCCCATATGCTATGAGTTACTGTGGGGTCCAAGGGCACTC 3909

Qy   4693 GCTGAAACCAGCTATGTGAAAGTCCTTGAGTATGTGATCAAGGTCAGTGCAAGAGTTCGC 4752

Db   3910 GCTGCT----------TGAAAGTACTGGAGCACGTGGTCAGGGTCAATGCAAGAGTTCTC 3959

Qy   4753 TTTTTCTTCCCATCCCTGCGTGAAGCAGCTTTGAGAGAGGAGGAAGAGGGAGTCTGAGCA 4812

Db   3960 ATTTCCTACCCATCCCTGCGTGAAGCAGCTTTGAGAGAGGAGGAAGAGGGAGTCTGAGCA 4019

Qy   4813 TGAGTTGCAGCCAAGGCCAGTGGGAGGGGGACTGGGCCAGTGCACCTTCCAGGGCCGCGT 4872

Db   4020 TGAGCTGCAGCCAGGGCCACTGCGAGGGGGGCTGGGCCAGTGCACCTTCCAGGGCTCCGT 4079

Qy   4873 CCAGCAGCTTCCCCTGCC-TCGTGTGACATGAGGCCCATTCTTCACTC--TGAAGAGAGC 4929

Db   4080 CCAGTAGTTTCCCCTGCCTTAATGTGACATGAGGCCCATTCTTCTCTCTTTGAAGAGAGC 4139

Qy   4930 GGTCAGTGTTCTCAGTAGTAGGTTTCTGTTCTATTGGGTGACTTGGAGATTTATCTTTGT 4989

Db   4140 AGTCAACATTCTTAGTAGTGGGTTTCTGTTCTATTGGATGACTTTGAGATTTGTCTTTGT 4199

Qy   4990 TCTCTTTTGGAATTGTTCAAATGTTTTTTTTTAAGGGATGGTTGAATGAACTTCAGCATC 5049

Db   4200 TTCCTTTTGGAATTGTTCAAATG-TTTCTTTTAATGGGTGGTTGAATGAACTTCAGCATT 4258

Qy   5050 CAAGTTTATGAATGACAGCAGTCACAC--AGTTCTGTGTATATAGTTTAAGGGTAAGAGT 5107

Db   4259 CAAATTTATGAATGACAGTAGTCACACATAGTGCTGTTTATATAGTTTAGGAGTAAGAGT 4318

Qy   5108 CTTGTGTTTTATTCAGATTGGGAAATCCATTCTATTTTGTGAATTGGG--ATAATAACAG 5165

Db   4319 CTTGTTTTTTATTCAGATTGGGAAATCCATTCCATTTTGTGAATTGGGACATAGTTACAG 4378

Qy   5166 CAGTGGAATAAGTACTTA-GAAATGTGAAAAATGAGCAGTAAAATAGATGAGATAAAGAA 5224

Db   4379 CAGTGGAATAAGTATTCATTTAGAAATGTGAATGAGCAGTAAAACTGATGACA------- 4431

Qy   5225 CTAAAGAAATTAAGAGATAGTCAATTCTTGCCTTATACCTCAGTCTATTCTGTAAAATTT 5284

Db   4432 -TAAAGAAATTAAAAGATATTTAATTCTTGCTTATACTCAGTCTATTCGGTAAAATTTTT 4490

Qy   5285 TTAAAGATATATGCATACCTGGATTTCCTTGGCTTCTTTGAGAATGTAAGAGAAATTAAA 5344

Db   4491 TTTAAAAAATGTGCATACCTGGATTTCCTTGGCTTCTTTGAGAATGTAAGACAAATTAAA 4550

Qy   5345 TCTGAATAAAGAATTCTTCCTGTTCACTGGCTCTTTTCTTCTCCATGCACTGAGCATCTG 5404

Db   4551 TCTGAATAAATCATTCTCCCTGTTCACTGGCTCATTTATTCTCTATGCACTGAGCATTTG 4610

Qy   5405 CTTTTTGGAAGGCCCTGGGTTAGTAGTGGAGATGCTAAGGTAAGCCAGACTCATACCCAC 5464

Db   4611 CTCTGTGGAAGGCCCTGGGTTAATAGTGGAGATGCTAAGGTAAGCCAGACTCACCCCTAC 4670

Qy   5465 CCATAGGGTCGTAGAGTCTAGGAGCTGCAGTCACGTAATCGAGGTGGCAAGATGTCCTCT 5524

Db   4671 CCACAGGGTAGTAAAGTCTAGGAGCAGCAGTCATATAATTAAGGTGGAGAGATGCCCTCT 4730

Qy   5525 AA 5526

Db   4731 AA 4732
```

RESULT 12
I36922
LOCUS I36922 2419 bp DNA PAT 13-MAY-1997
DEFINITION Sequence 7 from patent US 5612201.
ACCESSION I36922
VERSION I36922.1 GI:2084882
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 2419)
AUTHORS De Plaen,E., Boon-Falleur,T., Lethe ,B., Szikora,J., De Smet,C. and Chomez,P.
TITLE Isolated nucleic acid molecules useful in determining expression of a tumor rejection antigen precursor
JOURNAL Patent: US 5612201-A 7 18-MAR-1997;
FEATURES Location/Qualifiers
source 1. .2419
/organism="unknown"
BASE COUNT 562 a 581 c 677 g 599 t
ORIGIN

Query Match 42.6%; Score 2419; DB 5; Length 2419;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2419; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy   3256 GGATCCAGGCCCTGCCAGGAAAAATATAAGGGCCCTGCGTGAGAACAGAGGGGGTCATCC 3315

Db      1 GGATCCAGGCCCTGCCAGGAAAAATATAAGGGCCCTGCGTGAGAACAGAGGGGGTCATCC 60

Qy   3316 ACTGCATGAGAGTGGGGATGTCACAGAGTCCAGCCCACCCTCCTGGTAGCACTGAGAAGC 3375

Db     61 ACTGCATGAGAGTGGGGATGTCACAGAGTCCAGCCCACCCTCCTGGTAGCACTGAGAAGC 120

Qy   3376 CAGGGCTGTGCTTGCGGTCTGCACCCTGAGGGCCCGTGGATTCCTCTTCCTGGAGCTCCA 3435

Db    121 CAGGGCTGTGCTTGCGGTCTGCACCCTGAGGGCCCGTGGATTCCTCTTCCTGGAGCTCCA 180

Qy   3436 GGAACCAGGCAGTGAGGCCTTGGTCTGAGACAGTATCCTCAGGTCACAGAGCAGAGGATG 3495

Db    181 GGAACCAGGCAGTGAGGCCTTGGTCTGAGACAGTATCCTCAGGTCACAGAGCAGAGGATG 240

Qy   3496 CACAGGGTGTGCCAGCAGTGAATGTTTGCCCTGAATGCACACCAAGGGCCCCACCTGCCA 3555

Db    241 CACAGGGTGTGCCAGCAGTGAATGTTTGCCCTGAATGCACACCAAGGGCCCCACCTGCCA 300

Qy   3556 CAGGACACATAGGACTCCACAGAGTCTGGCCTCACCTCCCTACTGTCAGTCCTGTAGAAT 3615

Db    301 CAGGACACATAGGACTCCACAGAGTCTGGCCTCACCTCCCTACTGTCAGTCCTGTAGAAT 360
```

```
Qy   3616 CGACCTCTGCTGGCCGGCTGTACCCTGAGTACCCTCTCACTTCCTCCTTCAGGTTTTCAG 3675
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CGACCTCTGCTGGCCGGCTGTACCCTGAGTACCCTCTCACTTCCTCCTTCAGGTTTTCAG 420

Qy   3676 GGGACAGGCCAACCCAGAGGACAGGATTCCCTGGAGGCCACAGAGGAGCACCAAGGAGAA 3735
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GGGACAGGCCAACCCAGAGGACAGGATTCCCTGGAGGCCACAGAGGAGCACCAAGGAGAA 480

Qy   3736 GATCTGTAAGTAGGCCTTTGTTAGAGTCTCCAAGGTTCAGTTCTCAGCTGAGGCCTCTCA 3795
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GATCTGTAAGTAGGCCTTTGTTAGAGTCTCCAAGGTTCAGTTCTCAGCTGAGGCCTCTCA 540

Qy   3796 CACACTCCCTCTCTCCCCAGGCCTGTGGGTCTTCATTGCCCAGCTCCTGCCCACACTCCT 3855
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CACACTCCCTCTCTCCCCAGGCCTGTGGGTCTTCATTGCCCAGCTCCTGCCCACACTCCT 600

Qy   3856 GCCTGCTGCCCTGACGAGAGTCATCATGTCTCTTGAGCAGAGGAGTCTGCACTGCAAGCC 3915
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GCCTGCTGCCCTGACGAGAGTCATCATGTCTCTTGAGCAGAGGAGTCTGCACTGCAAGCC 660

Qy   3916 TGAGGAAGCCCTTGAGGCCCAACAAGAGGCCCTGGGCCTGGTGTGTGTGCAGGCTGCCAC 3975
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TGAGGAAGCCCTTGAGGCCCAACAAGAGGCCCTGGGCCTGGTGTGTGTGCAGGCTGCCAC 720

Qy   3976 CTCCTCCTCCTCTCCTCTGGTCCTGGGCACCCTGGAGGAGGTGCCCACTGCTGGGTCAAC 4035
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CTCCTCCTCCTCTCCTCTGGTCCTGGGCACCCTGGAGGAGGTGCCCACTGCTGGGTCAAC 780

Qy   4036 AGATCCTCCCCAGAGTCCTCAGGGAGCCTCCGCCTTTCCCACTACCATCAACTTCACTCG 4095
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 AGATCCTCCCCAGAGTCCTCAGGGAGCCTCCGCCTTTCCCACTACCATCAACTTCACTCG 840

Qy   4096 ACAGAGGCAACCCAGTGAGGGTTCCAGCAGCCGTGAAGAGGAGGGGCCAAGCACCTCTTG 4155
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 ACAGAGGCAACCCAGTGAGGGTTCCAGCAGCCGTGAAGAGGAGGGGCCAAGCACCTCTTG 900

Qy   4156 TATCCTGGAGTCCTTGTTCCGAGCAGTAATCACTAAGAAGGTGGCTGATTTGGTTGGTTT 4215
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TATCCTGGAGTCCTTGTTCCGAGCAGTAATCACTAAGAAGGTGGCTGATTTGGTTGGTTT 960

Qy   4216 TCTGCTCCTCAAATATCGAGCCAGGGAGCCAGTCACAAAGGCAGAAATGCTGGAGAGTGT 4275
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TCTGCTCCTCAAATATCGAGCCAGGGAGCCAGTCACAAAGGCAGAAATGCTGGAGAGTGT 1020

Qy   4276 CATCAAAAATTACAAGCACTGTTTTCCTGAGATCTTCGGCAAAGCCTCTGAGTCCTTGCA 4335
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CATCAAAAATTACAAGCACTGTTTTCCTGAGATCTTCGGCAAAGCCTCTGAGTCCTTGCA 1080

Qy   4336 GCTGGTCTTTGGCATTGACGTGAAGGAAGCAGACCCCACCGGCCACTCCTATGTCCTTGT 4395
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 GCTGGTCTTTGGCATTGACGTGAAGGAAGCAGACCCCACCGGCCACTCCTATGTCCTTGT 1140

Qy   4396 CACCTGCCTAGGTCTCTCCTATGATGGCCTGCTGGGTGATAATCAGATCATGCCCAAGAC 4455
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CACCTGCCTAGGTCTCTCCTATGATGGCCTGCTGGGTGATAATCAGATCATGCCCAAGAC 1200

Qy   4456 AGGCTTCCTGATAATTGTCCTGGTCATGATTGCAATGGAGGGCGGCCATGCTCCTGAGGA 4515
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 AGGCTTCCTGATAATTGTCCTGGTCATGATTGCAATGGAGGGCGGCCATGCTCCTGAGGA 1260

Qy   4516 GGAAATCTGGGAGGAGCTGAGTGTGATGGAGGTGTATGATGGGAGGGAGCACAGTGCCTA 4575
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GGAAATCTGGGAGGAGCTGAGTGTGATGGAGGTGTATGATGGGAGGGAGCACAGTGCCTA 1320

Qy   4576 TGGGGAGCCCAGGAAGCTGCTCACCCAAGATTTGGTGCAGGAAAAGTACCTGGAGTACGG 4635
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 TGGGGAGCCCAGGAAGCTGCTCACCCAAGATTTGGTGCAGGAAAAGTACCTGGAGTACGG 1380

Qy   4636 CAGGTGCCGGACAGTGATCCCGCACGCTATGAGTTCCTGTGGGGTCCAAGGGCCCTCGCT 4695
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 CAGGTGCCGGACAGTGATCCCGCACGCTATGAGTTCCTGTGGGGTCCAAGGGCCCTCGCT 1440

Qy   4696 GAAACCAGCTATGTGAAAGTCCTTGAGTATGTGATCAAGGTCAGTGCAAGAGTTCGCTTT 4755
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 GAAACCAGCTATGTGAAAGTCCTTGAGTATGTGATCAAGGTCAGTGCAAGAGTTCGCTTT 1500

Qy   4756 TTCTTCCCATCCCTGCGTGAAGCAGCTTTGAGAGAGGAGGAAGAGGGAGTCTGAGCATGA 4815
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 TTCTTCCCATCCCTGCGTGAAGCAGCTTTGAGAGAGGAGGAAGAGGGAGTCTGAGCATGA 1560

Qy   4816 GTTGCAGCCAAGGCCAGTGGGAGGGGGACTGGGCCAGTGCACCTTCCAGGGCCGCGTCCA 4875
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 GTTGCAGCCAAGGCCAGTGGGAGGGGGACTGGGCCAGTGCACCTTCCAGGGCCGCGTCCA 1620

Qy   4876 GCAGCTTCCCCTGCCTCGTGTGACATGAGGCCCATTCTTCACTCTGAAGAGAGCGGTCAG 4935
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 GCAGCTTCCCCTGCCTCGTGTGACATGAGGCCCATTCTTCACTCTGAAGAGAGCGGTCAG 1680

Qy   4936 TGTTCTCAGTAGTAGGTTTCTGTTCTATTGGGTGACTTGGAGATTTATCTTTGTTCTCTT 4995
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 TGTTCTCAGTAGTAGGTTTCTGTTCTATTGGGTGACTTGGAGATTTATCTTTGTTCTCTT 1740

Qy   4996 TTGGAATTGTTCAAATGTTTTTTTTTAAGGGATGGTTGAATGAACTTCAGCATCCAAGTT 5055
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 TTGGAATTGTTCAAATGTTTTTTTTTAAGGGATGGTTGAATGAACTTCAGCATCCAAGTT 1800

Qy   5056 TATGAATGACAGCAGTCACACAGTTCTGTGTATATAGTTTAAGGGTAAGAGTCTTGTGTT 5115
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 TATGAATGACAGCAGTCACACAGTTCTGTGTATATAGTTTAAGGGTAAGAGTCTTGTGTT 1860

Qy   5116 TTATTCAGATTGGGAAATCCATTCTATTTTGTGAATTGGGATAATAACAGCAGTGGAATA 5175
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 TTATTCAGATTGGGAAATCCATTCTATTTTGTGAATTGGGATAATAACAGCAGTGGAATA 1920

Qy   5176 AGTACTTAGAAATGTGAAAAATGAGCAGTAAAATAGATGAGATAAAGAACTAAAGAAATT 5235
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 AGTACTTAGAAATGTGAAAAATGAGCAGTAAAATAGATGAGATAAAGAACTAAAGAAATT 1980

Qy   5236 AAGAGATAGTCAATTCTTGCCTTATACCTCAGTCTATTCTGTAAAATTTTTAAAGATATA 5295
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 AAGAGATAGTCAATTCTTGCCTTATACCTCAGTCTATTCTGTAAAATTTTTAAAGATATA 2040

Qy   5296 TGCATACCTGGATTTCCTTGGCTTCTTTGAGAATGTAAGAGAAATTAAATCTGAATAAAG 5355
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 TGCATACCTGGATTTCCTTGGCTTCTTTGAGAATGTAAGAGAAATTAAATCTGAATAAAG 2100

Qy   5356 AATTCTTCCTGTTCACTGGCTCTTTTCTTCTCCATGCACTGAGCATCTGCTTTTTGGAAG 5415
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 AATTCTTCCTGTTCACTGGCTCTTTTCTTCTCCATGCACTGAGCATCTGCTTTTTGGAAG 2160

Qy   5416 GCCCTGGGTTAGTAGTGGAGATGCTAAGGTAAGCCAGACTCATACCCACCCATAGGGTCG 5475
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 GCCCTGGGTTAGTAGTGGAGATGCTAAGGTAAGCCAGACTCATACCCACCCATAGGGTCG 2220

Qy   5476 TAGAGTCTAGGAGCTGCAGTCACGTAATCGAGGTGGCAAGATGTCCTCTAAAGATGTAGG 5535
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 TAGAGTCTAGGAGCTGCAGTCACGTAATCGAGGTGGCAAGATGTCCTCTAAAGATGTAGG 2280

Qy   5536 GAAAAGTGAGAGAGGGGTGAGGGTGTGGGGCTCCGGGTGAGAGTGGTGGAGTGTCAATGC 5595
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 GAAAAGTGAGAGAGGGGTGAGGGTGTGGGGCTCCGGGTGAGAGTGGTGGAGTGTCAATGC 2340

Qy   5596 CCTGAGCTGGGGCATTTTGGGCTTTGGGAAACTGCAGTTCCTTCTGGGGGAGCTGATTGT 5655
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 CCTGAGCTGGGGCATTTTGGGCTTTGGGAAACTGCAGTTCCTTCTGGGGGAGCTGATTGT 2400

Qy   5656 AATGATCTTGGGTGGATCC 5674
          ||||||||||||||||||
Db   2401 AATGATCTTGGGTGGATCC 2419
```

RESULT 13
AR007331
LOCUS AR007331 2420 bp DNA PAT 04-DEC-1998
DEFINITION Sequence 23 from patent US 5750395.

```
ACCESSION   AR007331
VERSION     AR007331.1  GI:3966815
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2420)
  AUTHORS   Fikes,J.D., Livingston,B.D., Sette,A.D. and Sidney,J.C.
  TITLE     DNA encoding MAGE-1 C-terminal cytotoxic t lymphocyte immunogenic
            peptides
  JOURNAL   Patent: US 5750395-A 23 12-MAY-1998;
FEATURES             Location/Qualifiers
     source          1. .2420
                     /organism="unknown"
BASE COUNT      562 a    582 c    677 g    599 t
ORIGIN
```

Query Match 42.4%; Score 2408; DB 5; Length 2420;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2419; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy   3256 GGATCCAGGCCCTGCCAGGAAAAATATAAGGGCCCTGCGTGAGAACAGAGGGGGTCATCC 3315
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGATCCAGGCCCTGCCAGGAAAAATATAAGGGCCCTGCGTGAGAACAGAGGGGGTCATCC 60

Qy   3316 ACTGCATGAGAGTGGGGATGTCACAGAGTCCAGCCCACCCTCCTGGTAGCACTGAGAAGC 3375
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ACTGCATGAGAGTGGGGATGTCACAGAGTCCAGCCCACCCTCCTGGTAGCACTGAGAAGC 120

Qy   3376 CAGGGCTGTGCTTGCGGTCTGCACCCTGAGGGCCCGTGGATTCCTCTTCCTGGAGCTCCA 3435
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CAGGGCTGTGCTTGCGGTCTGCACCCTGAGGGCCCGTGGATTCCTCTTCCTGGAGCTCCA 180

Qy   3436 GGAACCAGGCAGTGAGGCCTTGGTCTGAGACAGTATCCTCAGGTCACAGAGCAGAGGATG 3495
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GGAACCAGGCAGTGAGGCCTTGGTCTGAGACAGTATCCTCAGGTCACAGAGCAGAGGATG 240

Qy   3496 CACAGGGTGTGCCAGCAGTGAATGTTTGCCCTGAATGCACACCAAGGGCCCCACCTGCCA 3555
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CACAGGGTGTGCCAGCAGTGAATGTTTGCCCTGAATGCACACCAAGGGCCCCACCTGCCA 300

Qy   3556 CAGGACACATAGGACTCCACAGAGTCTGGCCTCACCTCCCTACTGTCAGTCCTGTAGAAT 3615
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CAGGACACATAGGACTCCACAGAGTCTGGCCTCACCTCCCTACTGTCAGTCCTGTAGAAT 360

Qy   3616 CGACCTCTGCTGGCCGGCTGTACCCTGAGTACCCTCTCACTTCCTCCTTCAGGTTTTCAG 3675
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CGACCTCTGCTGGCCGGCTGTACCCTGAGTACCCTCTCACTTCCTCCTTCAGGTTTTCAG 420

Qy   3676 GGGACAGGCCAACCCAGAGGACAGGATTCCCTGGAGGCCACAGAGGAGCACCAAGGAGAA 3735
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GGGACAGGCCAACCCAGAGGACAGGATTCCCTGGAGGCCACAGAGGAGCACCAAGGAGAA 480

Qy   3736 GATCTGTAAGTAGGCCTTTGTTAGAGTCTCCAAGGTTCAGTTCTCAGCTGAGGCCTCTCA 3795
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GATCTGTAAGTAGGCCTTTGTTAGAGTCTCCAAGGTTCAGTTCTCAGCTGAGGCCTCTCA 540

Qy   3796 CACACTCCCTCTCTCCCCAGGCCTGTGGGTCTTCATTGCCCAGCTCCTGCCCACACTCCT 3855
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CACACTCCCTCTCTCCCCAGGCCTGTGGGTCTTCATTGCCCAGCTCCTGCCCACACTCCT 600

Qy   3856 GCCTGCTGCCCCTGACGAGAGTCATCATGTCTCTTGAGCAGAGGAGTCTGCACTGCAAGCC 3915
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GCCTGCTGCCCCTGACGAGAGTCATCATGTCTCTTGAGCAGAGGAGTCTGCACTGCAAGCC 660

Qy   3916 TGAGGAAGCCCTTGAGGCCCAACAAGAGGCCCTGGGCCTGGTGTGTGTGCAGGCTGCCAC 3975
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TGAGGAAGCCCTTGAGGCCCAACAAGAGGCCCTGGGCCTGGTGTGTGTGCAGGCTGCCAC 720

Qy   3976 CTCCTCCTCCTCTCCTCTGGTCCTGGGCACCCTGGAGGAGGTGCCCACTGCTGGGTCAAC 4035
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CTCCTCCTCCTCTCCTCTGGTCCTGGGCACCCTGGAGGAGGTGCCCACTGCTGGGTCAAC 780

Qy   4036 AGATCCTCCCCAGAGTCCTCAGGGAGCCTCCGCCTTTCCCACTACCATCAACTTCACTCG 4095
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 AGATCCTCCCCAGAGTCCTCAGGGAGCCTCCGCCTTTCCCACTACCATCAACTTCACTCG 840

Qy   4096 ACAGAGGCAACCCAGTGAGGGTTCCAGCAGCCGTGAAGAGGAGGGGCCAAGCACCTCTTG 4155
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 ACAGAGGCAACCCAGTGAGGGTTCCAGCAGCCGTGAAGAGGAGGGGCCAAGCACCTCTTG 900

Qy   4156 TATCCTGGAGTCCTTGTTCCGAGCAGTAATCACTAAGAAGGTGGCTGATTTGGTTGGTTT 4215
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TATCCTGGAGTCCTTGTTCCGAGCAGTAATCACTAAGAAGGTGGCTGATTTGGTTGGTTT 960

Qy   4216 TCTGCTCCTCAAATATCGAGCCAGGGAGCCAGTCACAAAGGCAGAAATGCTGGAGAGTGT 4275
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TCTGCTCCTCAAATATCGAGCCAGGGAGCCAGTCACAAAGGCAGAAATGCTGGAGAGTGT 1020

Qy   4276 CATCAAAAATTACAAGCACTGTTTTCCTGAGATCTTCGGCAAAGCCTCTGAGTCCTTGCA 4335
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CATCAAAAATTACAAGCACTGTTTTCCTGAGATCTTCGGCAAAGCCTCTGAGTCCTTGCA 1080

Qy   4336 GCTGGTCTTTGGCATTGACGTGAAGGAAGCAGACCCCACCGGCCACTCCTATGTCCTTGT 4395
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 GCTGGTCTTTGGCATTGACGTGAAGGAAGCAGACCCCACCGGCCACTCCTATGTCCTTGT 1140

Qy   4396 CACCTGCCTAGGTCTCTCCTATGATGGCCTGCTGGGTGATAATCAGATCATGCCCAAGAC 4455
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CACCTGCCTAGGTCTCTCCTATGATGGCCTGCTGGGTGATAATCAGATCATGCCCAAGAC 1200

Qy   4456 AGGCTTCCTGATAATTGTCCTGGTCATGATTGCAATGGAGGGCGGCCATGCTCCTGAGGA 4515
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 AGGCTTCCTGATAATTGTCCTGGTCATGATTGCAATGGAGGGCGGCCATGCTCCTGAGGA 1260

Qy   4516 GGAAATCTGGGAGGAGCTGAGTGTGATGGAGGTGTATGATGGGAGGGAGCACAGTGCCTA 4575
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GGAAATCTGGGAGGAGCTGAGTGTGATGGAGGTGTATGATGGGAGGGAGCACAGTGCCTA 1320

Qy   4576 TGGGGAGCCCAGGAAGCTGCTCACCCAAGATTTGGTGCAGGAAAAGTACCTGGAGTA-CG 4634
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db   1321 TGGGGAGCCCAGGAAGCTGCTCACCCAAGATTTGGTGCAGGAAAAGTACCTGGAGTACCG 1380

Qy   4635 GCAGGTGCCGGACAGTGATCCCGCACGCTATGAGTTCCTGTGGGGTCCAAGGGCCCTCGC 4694
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 GCAGGTGCCGGACAGTGATCCCGCACGCTATGAGTTCCTGTGGGGTCCAAGGGCCCTCGC 1440

Qy   4695 TGAAACCAGCTATGTGAAAGTCCTTGAGTATGTGATCAAGGTCAGTGCAAGAGTTCGCTT 4754
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 TGAAACCAGCTATGTGAAAGTCCTTGAGTATGTGATCAAGGTCAGTGCAAGAGTTCGCTT 1500

Qy   4755 TTTCTTCCCATCCCTGCGTGAAGCAGCTTTGAGAGAGGAGGAAGAGGGAGTCTGAGCATG 4814
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 TTTCTTCCCATCCCTGCGTGAAGCAGCTTTGAGAGAGGAGGAAGAGGGAGTCTGAGCATG 1560

Qy   4815 AGTTGCAGCCAAGGCCAGTGGGAGGGGGACTGGGCCAGTGCACCTTCCAGGGCCGCGTCC 4874
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 AGTTGCAGCCAAGGCCAGTGGGAGGGGGACTGGGCCAGTGCACCTTCCAGGGCCGCGTCC 1620

Qy   4875 AGCAGCTTCCCCTGCCTCGTGTGACATGAGGCCCATTCTTCACTCTGAAGAGAGCGGTCA 4934
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 AGCAGCTTCCCCTGCCTCGTGTGACATGAGGCCCATTCTTCACTCTGAAGAGAGCGGTCA 1680

Qy   4935 GTGTTCTCAGTAGTAGGTTTCTGTTCTATTGGGTGACTTGGAGATTTATCTTTGTTCTCT 4994
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 GTGTTCTCAGTAGTAGGTTTCTGTTCTATTGGGTGACTTGGAGATTTATCTTTGTTCTCT 1740

Qy   4995 TTTGGAATTGTTCAAATGTTTTTTTTTAAGGGATGGTTGAATGAACTTCAGCATCCAAGT 5054
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 TTTGGAATTGTTCAAATGTTTTTTTTTAAGGGATGGTTGAATGAACTTCAGCATCCAAGT 1800

Qy   5055 TTATGAATGACAGCAGTCACACAGTTCTGTGTATATAGTTTAAGGGTAAGAGTCTTGTGT 5114
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 TTATGAATGACAGCAGTCACACAGTTCTGTGTATATAGTTTAAGGGTAAGAGTCTTGTGT 1860
```

```
Qy   5115 TTTATTCAGATTGGGAAATCCATTCTATTTTGTGAATTGGGATAATAACAGCAGTGGAAT 5174
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 TTTATTCAGATTGGGAAATCCATTCTATTTTGTGAATTGGGATAATAACAGCAGTGGAAT 1920

Qy   5175 AAGTACTTAGAAATGTGAAAAATGAGCAGTAAAATAGATGAGATAAAGAACTAAAGAAAT 5234
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 AAGTACTTAGAAATGTGAAAAATGAGCAGTAAAATAGATGAGATAAAGAACTAAAGAAAT 1980

Qy   5235 TAAGAGATAGTCAATTCTTGCCTTATACCTCAGTCTATTCTGTAAAATTTTTAAAGATAT 5294
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 TAAGAGATAGTCAATTCTTGCCTTATACCTCAGTCTATTCTGTAAAATTTTTAAAGATAT 2040

Qy   5295 ATGCATACCTGGATTTCCTTGGCTTCTTTGAGAATGTAAGAGAAATTAAATCTGAATAAA 5354
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 ATGCATACCTGGATTTCCTTGGCTTCTTTGAGAATGTAAGAGAAATTAAATCTGAATAAA 2100

Qy   5355 GAATTCTTCCTGTTCACTGGCTCTTTTCTTCTCCATGCACTGAGCATCTGCTTTTTGGAA 5414
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 GAATTCTTCCTGTTCACTGGCTCTTTTCTTCTCCATGCACTGAGCATCTGCTTTTTGGAA 2160

Qy   5415 GGCCCTGGGTTAGTAGTGGAGATGCTAAGGTAAGCCAGACTCATACCCACCCATAGGGTC 5474
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 GGCCCTGGGTTAGTAGTGGAGATGCTAAGGTAAGCCAGACTCATACCCACCCATAGGGTC 2220

Qy   5475 GTAGAGTCTAGGAGCTGCAGTCACGTAATCGAGGTGGCAAGATGTCCTCTAAAGATGTAG 5534
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 GTAGAGTCTAGGAGCTGCAGTCACGTAATCGAGGTGGCAAGATGTCCTCTAAAGATGTAG 2280

Qy   5535 GGAAAAGTGAGAGAGGGGTGAGGGTGTGGGGCTCCGGGTGAGAGTGGTGGAGTGTCAATG 5594
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 GGAAAAGTGAGAGAGGGGTGAGGGTGTGGGGCTCCGGGTGAGAGTGGTGGAGTGTCAATG 2340

Qy   5595 CCCTGAGCTGGGGCATTTTGGGCTTTGGGAAACTGCAGTTCCTTCTGGGGGAGCTGATTG 5654
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 CCCTGAGCTGGGGCATTTTGGGCTTTGGGAAACTGCAGTTCCTTCTGGGGGAGCTGATTG 2400

Qy   5655 TAATGATCTTGGGTGGATCC 5674
          ||||||||||||||||||||
Db   2401 TAATGATCTTGGGTGGATCC 2420
```

```
RESULT  14
HUMMAG1A
LOCUS       HUMMAG1A     2420 bp    DNA             PRI       15-NOV-1993
DEFINITION  Human antigen (MAGE-1) gene, complete cds.
ACCESSION   M77481
VERSION     M77481.1  GI:416114
KEYWORDS    antigen.
SOURCE      Homo sapiens (individual_isolate patient MZ2) melanoma abdominal
            metastasis of melanoma DNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 785 to 1286)
  AUTHORS   van der Bruggen,P., Traversari,C., Chomez,P., Lurquin,C., De
            Plaen,E., Van den Eynde,B., Knuth,A. and Boon,T.
  TITLE     A gene encoding an antigen recognized by cytolytic T lymphocytes on
            a human melanoma
  JOURNAL   Science 254, 1643-1647 (1991)
  MEDLINE   92086861
REFERENCE   2  (bases 1 to 2420)
  AUTHORS   van der Bruggen P.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-FEB-1992) Pierre van der Bruggen, Ludwig Institute
            for Cancer Research, Brussels Branch, Avenue Hippocrate, 74, UCL
            7459, Brussels, B-1200, Belgium
COMMENT     On Nov 15, 1993 this sequence version replaced gi:187294.
FEATURES             Location/Qualifiers
     source          1. .2420
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="X"
                     /sex="female"
                     /cell_line="MZ2-MEL.43"
     intron          1. .412
                     /partial
                     /gene="MAGE-1"
                     /number=1
     exon            413. .485
                     /gene="MAGE-1"
                     /number=2
     mRNA            join(413. .485,561. .2111)
                     /partial
                     /gene="MAGE-1"
     intron          486. .560
                     /gene="MAGE-1"
                     /number=2
     exon            561. .2111
                     /gene="MAGE-1"
                     /number=3
     gene            626. .1555
                     /gene="MAGE-1"
     CDS             626. .1555
                     /gene="MAGE-1"
                     /codon_start=1
                     /product="MAGE-1"
                     /protein_id="AAA03229.1"
                     /db_xref="GI:416115"
                     /translation="MSLEQRSLHCKPEEALEAQQEALGLVCVQAATSSSSPLVLGTLE
                     EVPTAGSTDPPQSPQGASAFPTTINFTRQRQPSEGSSSREEEGPSTSCILESLFRAVI
                     TKKVADLVGFLLLKYRAREPVTKAEMLESVIKNYKHCFPEIFGKASESLQLVFGIDVK
                     EADPTGHSYVLVTCLGLSYDGLLGDNQIMPKTGFLIIVLVMIAMEGGHAPEEEIWEEL
                     SVMEVYDGREHSAYGEPRKLLTQDLVQEKYLEYRQVPDSDPARYEFLWGPRALAETSY
                     VKVLEYVIKVSARVRFFFPSLREAALREEEEGV"
     polyA_signal    2095. .2100
                     /gene="MAGE-1"
     intron          2112. .2420
                     /partial
                     /gene="MAGE-1"
                     /number=3
BASE COUNT      562 a    582 c    677 g    599 t
ORIGIN
```

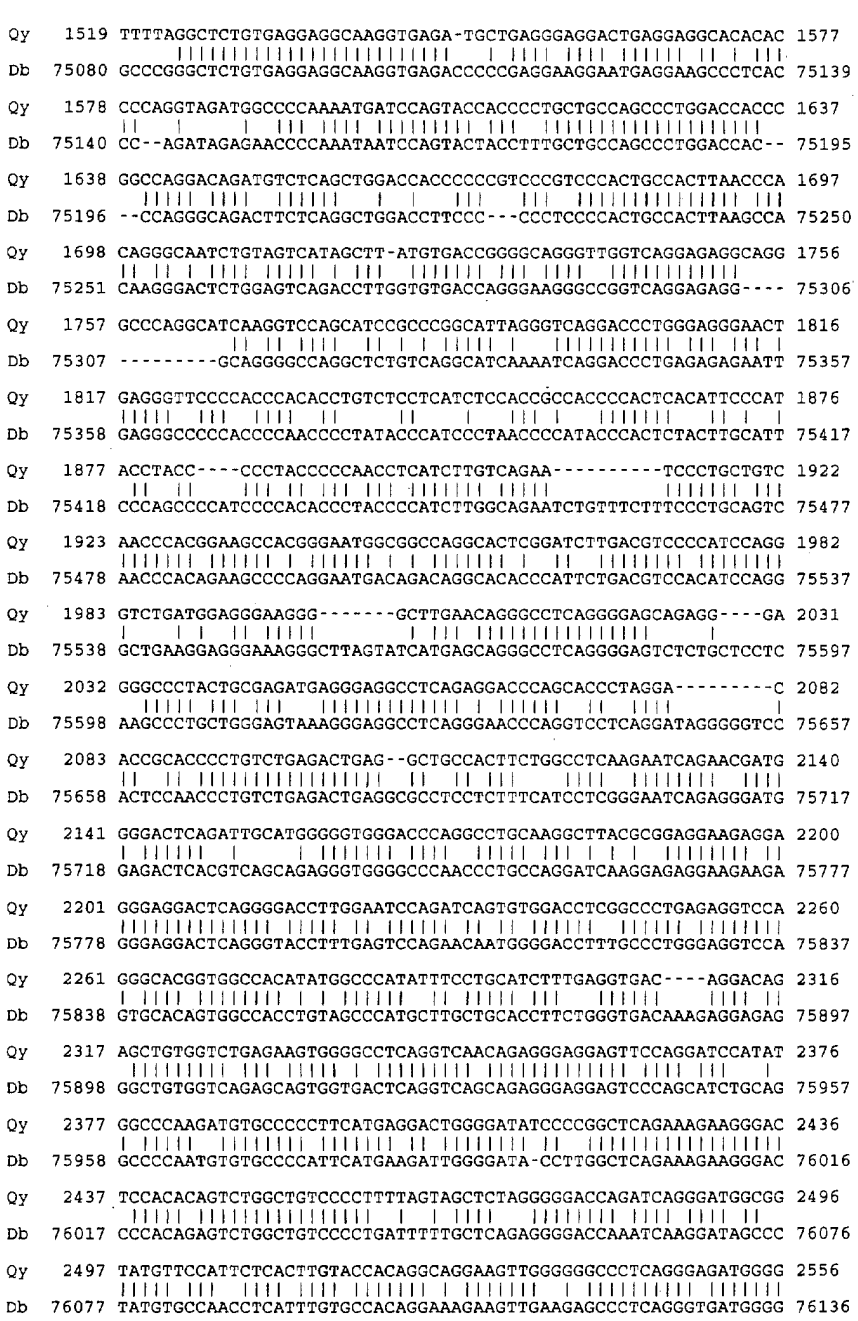
```
  Query Match           42.4%;  Score 2408;  DB 9;  Length 2420;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 2419;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy   3256 GGATCCAGGCCCTGCCAGGAAAAATATAAGGGCCCTGCGTGAGAACAGAGGGGGTCATCC 3315
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGATCCAGGCCCTGCCAGGAAAAATATAAGGGCCCTGCGTGAGAACAGAGGGGGTCATCC 60

Qy   3316 ACTGCATGAGAGTGGGGATGTCACAGAGTCCAGCCCACCCTCCTGGTAGCACTGAGAAGC 3375
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ACTGCATGAGAGTGGGGATGTCACAGAGTCCAGCCCACCCTCCTGGTAGCACTGAGAAGC 120

Qy   3376 CAGGGCTGTGCTTGCGGTCTGCACCCTGAGGGCCCGTGGATTCCTCTTCCTGGAGCTCCA 3435
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CAGGGCTGTGCTTGCGGTCTGCACCCTGAGGGCCCGTGGATTCCTCTTCCTGGAGCTCCA 180

Qy   3436 GGAACCAGGCAGTGAGGCCTTGGTCTGAGACAGTATCCTCAGGTCACAGAGCAGAGGATG 3495
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GGAACCAGGCAGTGAGGCCTTGGTCTGAGACAGTATCCTCAGGTCACAGAGCAGAGGATG 240

Qy   3496 CACAGGGTGTGCCAGCAGTGAATGTTTGCCCTGAATGCACACCAAGGGCCCCACCTGCCA 3555
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CACAGGGTGTGCCAGCAGTGAATGTTTGCCCTGAATGCACACCAAGGGCCCCACCTGCCA 300

Qy   3556 CAGGACACATAGGACTCCACAGAGTCTGGCCTCACCTCCCTACTGTCAGTCCTGTAGAAT 3615
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CAGGACACATAGGACTCCACAGAGTCTGGCCTCACCTCCCTACTGTCAGTCCTGTAGAAT 360

Qy   3616 CGACCTCTGCTGGCCGGCTGTACCCTGAGTACCCTCTCACTTCCTCCTTCAGGTTTTCAG 3675
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CGACCTCTGCTGGCCGGCTGTACCCTGAGTACCCTCTCACTTCCTCCTTCAGGTTTTCAG 420
```

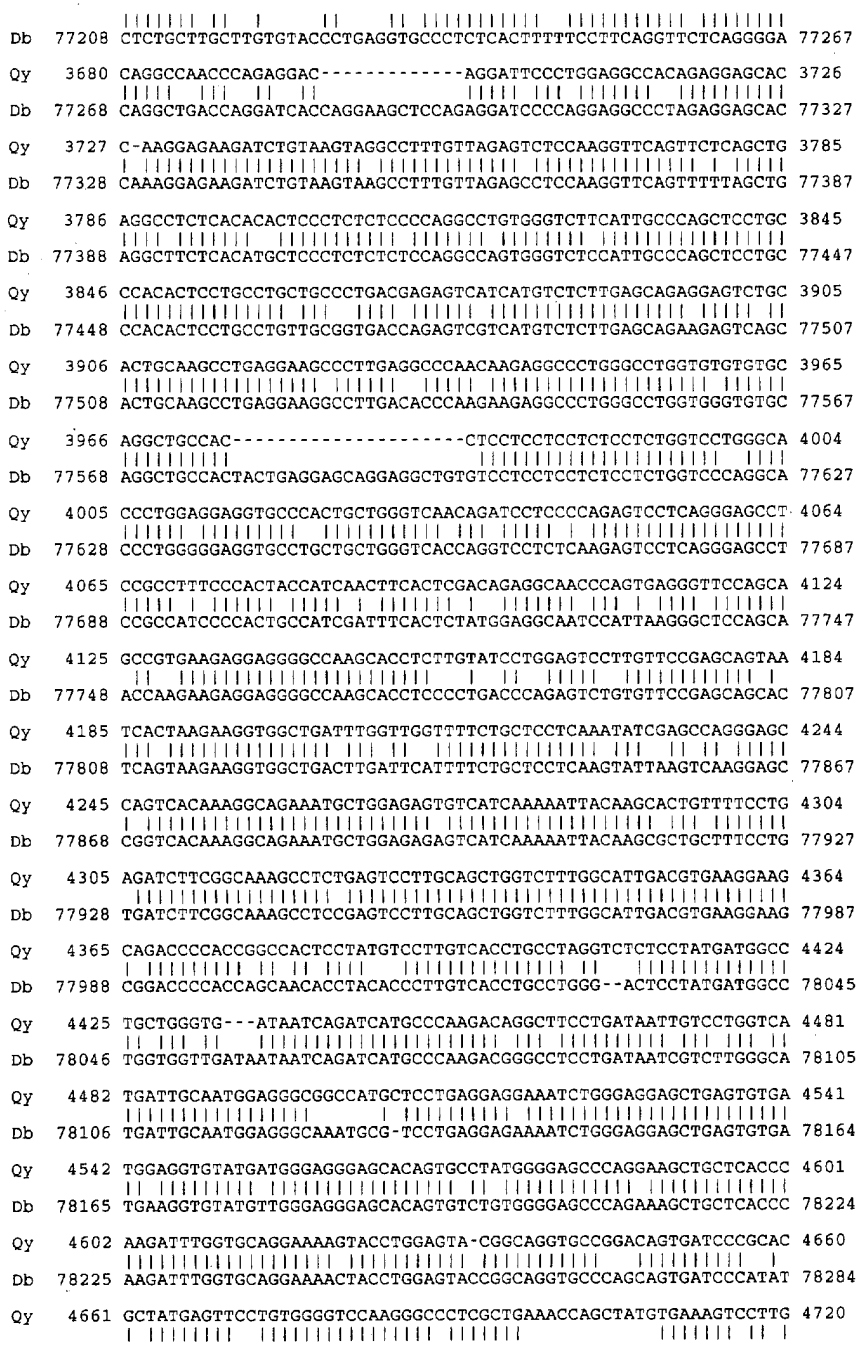
```
Qy   3676 GGGACAGGCCAACCCAGAGGACAGGATTCCCTGGAGGCCACAGAGGAGCACCAAGGAGAA 3735
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GGGACAGGCCAACCCAGAGGACAGGATTCCCTGGAGGCCACAGAGGAGCACCAAGGAGAA 480

Qy   3736 GATCTGTAAGTAGGCCTTTGTTAGAGTCTCCAAGGTTCAGTTCTCAGCTGAGGCCTCTCA 3795
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GATCTGTAAGTAGGCCTTTGTTAGAGTCTCCAAGGTTCAGTTCTCAGCTGAGGCCTCTCA 540

Qy   3796 CACACTCCCTCTCTCCCCAGGCCTGTGGGTCTTCATTGCCCAGCTCCTGCCCACACTCCT 3855
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CACACTCCCTCTCTCCCCAGGCCTGTGGGTCTTCATTGCCCAGCTCCTGCCCACACTCCT 600

Qy   3856 GCCTGCTGCCCTGACGAGAGTCATCATGTCTCTTGAGCAGAGGAGTCTGCACTGCAAGCC 3915
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GCCTGCTGCCCTGACGAGAGTCATCATGTCTCTTGAGCAGAGGAGTCTGCACTGCAAGCC 660

Qy   3916 TGAGGAAGCCCTTGAGGCCCAACAAGAGGCCCTGGGCCTGGTGTGTGTGCAGGCTGCCAC 3975
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TGAGGAAGCCCTTGAGGCCCAACAAGAGGCCCTGGGCCTGGTGTGTGTGCAGGCTGCCAC 720

Qy   3976 CTCCTCCTCCTCTCCTCTGGTCCTGGGCACCCTGGAGGAGGTGCCCACTGCTGGGTCAAC 4035
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CTCCTCCTCCTCTCCTCTGGTCCTGGGCACCCTGGAGGAGGTGCCCACTGCTGGGTCAAC 780

Qy   4036 AGATCCTCCCCAGAGTCCTCAGGGAGCCTCCGCCTTTCCCACTACCATCAACTTCACTCG 4095
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 AGATCCTCCCCAGAGTCCTCAGGGAGCCTCCGCCTTTCCCACTACCATCAACTTCACTCG 840

Qy   4096 ACAGAGGCAACCCAGTGAGGGTTCCAGCAGCCGTGAAGAGGAGGGGCCAAGCACCTCTTG 4155
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 ACAGAGGCAACCCAGTGAGGGTTCCAGCAGCCGTGAAGAGGAGGGGCCAAGCACCTCTTG 900

Qy   4156 TATCCTGGAGTCCTTGTTCCGAGCAGTAATCACTAAGAAGGTGGCTGATTTGGTTGGTTT 4215
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TATCCTGGAGTCCTTGTTCCGAGCAGTAATCACTAAGAAGGTGGCTGATTTGGTTGGTTT 960

Qy   4216 TCTGCTCCTCAAATATCGAGCCAGGGAGCCAGTCACAAAGGCAGAAATGCTGGAGAGTGT 4275
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TCTGCTCCTCAAATATCGAGCCAGGGAGCCAGTCACAAAGGCAGAAATGCTGGAGAGTGT 1020

Qy   4276 CATCAAAAATTACAAGCACTGTTTTCCTGAGATCTTCGGCAAAGCCTCTGAGTCCTTGCA 4335
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CATCAAAAATTACAAGCACTGTTTTCCTGAGATCTTCGGCAAAGCCTCTGAGTCCTTGCA 1080

Qy   4336 GCTGGTCTTTGGCATTGACGTGAAGGAAGCAGACCCCACCGGCCACTCCTATGTCCTTGT 4395
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 GCTGGTCTTTGGCATTGACGTGAAGGAAGCAGACCCCACCGGCCACTCCTATGTCCTTGT 1140

Qy   4396 CACCTGCCTAGGTCTCTCCTATGATGGCCTGCTGGGTGATAATCAGATCATGCCCAAGAC 4455
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CACCTGCCTAGGTCTCTCCTATGATGGCCTGCTGGGTGATAATCAGATCATGCCCAAGAC 1200

Qy   4456 AGGCTTCCTGATAATTGTCCTGGTCATGATTGCAATGGAGGGCGGCCATGCTCCTGAGGA 4515
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 AGGCTTCCTGATAATTGTCCTGGTCATGATTGCAATGGAGGGCGGCCATGCTCCTGAGGA 1260

Qy   4516 GGAAATCTGGGAGGAGCTGAGTGTGATGGAGGTGTATGATGGGAGGGAGCACAGTGCCTA 4575
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GGAAATCTGGGAGGAGCTGAGTGTGATGGAGGTGTATGATGGGAGGGAGCACAGTGCCTA 1320

Qy   4576 TGGGGAGCCCAGGAAGCTGCTCACCCAAGATTTGGTGCAGGAAAAGTACCTGGAGTA-CG 4634
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db   1321 TGGGGAGCCCAGGAAGCTGCTCACCCAAGATTTGGTGCAGGAAAAGTACCTGGAGTACCG 1380

Qy   4635 GCAGGTGCCGGACAGTGATCCCGCACGCTATGAGTTCCTGTGGGGTCCAAGGGCCCTCGC 4694
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 GCAGGTGCCGGACAGTGATCCCGCACGCTATGAGTTCCTGTGGGGTCCAAGGGCCCTCGC 1440

Qy   4695 TGAAACCAGCTATGTGAAAGTCCTTGAGTATGTGATCAAGGTCAGTGCAAGAGTTCGCTT 4754
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 TGAAACCAGCTATGTGAAAGTCCTTGAGTATGTGATCAAGGTCAGTGCAAGAGTTCGCTT 1500

Qy   4755 TTTCTTCCCATCCCTGCGTGAAGCAGCTTTGAGAGAGGAGGAAGAGGGAGTCTGAGCATG 4814
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 TTTCTTCCCATCCCTGCGTGAAGCAGCTTTGAGAGAGGAGGAAGAGGGAGTCTGAGCATG 1560

Qy   4815 AGTTGCAGCCAAGGCCAGTGGGAGGGGGACTGGGCCAGTGCACCTTCCAGGGCCGCGTCC 4874
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 AGTTGCAGCCAAGGCCAGTGGGAGGGGGACTGGGCCAGTGCACCTTCCAGGGCCGCGTCC 1620

Qy   4875 AGCAGCTTCCCCTGCCTCGTGTGACATGAGGCCCATTCTTCACTCTGAAGAGAGCGGTCA 4934
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 AGCAGCTTCCCCTGCCTCGTGTGACATGAGGCCCATTCTTCACTCTGAAGAGAGCGGTCA 1680

Qy   4935 GTGTTCTCAGTAGTAGGTTTCTGTTCTATTGGGTGACTTGGAGATTTATCTTTGTTCTCT 4994
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 GTGTTCTCAGTAGTAGGTTTCTGTTCTATTGGGTGACTTGGAGATTTATCTTTGTTCTCT 1740

Qy   4995 TTTGGAATTGTTCAAATGTTTTTTTTTAAGGGATGGTTGAATGAACTTCAGCATCCAAGT 5054
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 TTTGGAATTGTTCAAATGTTTTTTTTTAAGGGATGGTTGAATGAACTTCAGCATCCAAGT 1800

Qy   5055 TTATGAATGACAGCAGTCACACAGTTCTGTGTATATAGTTTAAGGGTAAGAGTCTTGTGT 5114
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 TTATGAATGACAGCAGTCACACAGTTCTGTGTATATAGTTTAAGGGTAAGAGTCTTGTGT 1860

Qy   5115 TTTATTCAGATTGGGAAATCCATTCTATTTTGTGAATTGGGATAATAACAGCAGTGGAAT 5174
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 TTTATTCAGATTGGGAAATCCATTCTATTTTGTGAATTGGGATAATAACAGCAGTGGAAT 1920

Qy   5175 AAGTACTTAGAAATGTGAAAAATGAGCAGTAAAATAGATGAGATAAAGAACTAAAGAAAT 5234
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 AAGTACTTAGAAATGTGAAAAATGAGCAGTAAAATAGATGAGATAAAGAACTAAAGAAAT 1980

Qy   5235 TAAGAGATAGTCAATTCTTGCCTTATACCTCAGTCTATTCTGTAAAATTTTTAAAGATAT 5294
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 TAAGAGATAGTCAATTCTTGCCTTATACCTCAGTCTATTCTGTAAAATTTTTAAAGATAT 2040

Qy   5295 ATGCATACCTGGATTTCCTTGGCTTCTTTGAGAATGTAAGAGAAATTAAATCTGAATAAA 5354
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 ATGCATACCTGGATTTCCTTGGCTTCTTTGAGAATGTAAGAGAAATTAAATCTGAATAAA 2100

Qy   5355 GAATTCTTCCTGTTCACTGGCTCTTTTCTTCTCCATGCACTGAGCATCTGCTTTTTGGAA 5414
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 GAATTCTTCCTGTTCACTGGCTCTTTTCTTCTCCATGCACTGAGCATCTGCTTTTTGGAA 2160

Qy   5415 GGCCCTGGGTTAGTAGTGGAGATGCTAAGGTAAGCCAGACTCATACCCACCCATAGGGTC 5474
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 GGCCCTGGGTTAGTAGTGGAGATGCTAAGGTAAGCCAGACTCATACCCACCCATAGGGTC 2220

Qy   5475 GTAGAGTCTAGGAGCTGCAGTCACGTAATCGAGGTGGCAAGATGTCCTCTAAAGATGTAG 5534
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 GTAGAGTCTAGGAGCTGCAGTCACGTAATCGAGGTGGCAAGATGTCCTCTAAAGATGTAG 2280

Qy   5535 GGAAAAGTGAGAGAGGGGTGAGGGTGTGGGGCTCCGGGTGAGAGTGGTGGAGTGTCAATG 5594
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 GGAAAAGTGAGAGAGGGGTGAGGGTGTGGGGCTCCGGGTGAGAGTGGTGGAGTGTCAATG 2340

Qy   5595 CCCTGAGCTGGGGCATTTTGGGCTTTGGGAAACTGCAGTTCCTTCTGGGGGAGCTGATTG 5654
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 CCCTGAGCTGGGGCATTTTGGGCTTTGGGAAACTGCAGTTCCTTCTGGGGGAGCTGATTG 2400

Qy   5655 TAATGATCTTGGGTGGATCC 5674
          ||||||||||||||||||||
Db   2401 TAATGATCTTGGGTGGATCC 2420
```

RESULT 15
AC009621
LOCUS AC009621 178515 bp DNA HTG 08-APR-2000
DEFINITION Homo sapiens chromosome 8 clone RP11-173J18 map 8, WORKING DRAFT SEQUENCE, 28 unordered pieces.
ACCESSION AC009621
VERSION AC009621.4 GI:7528051
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 178515)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 8, clone RP11-173J18
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 178515)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
            Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Brown,A.,
            Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A.,
            Cooke,P., DeArellano,K., Depayre,E., Devon,K., Dewar,K.,
            Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C.,
            Funke,R., Gage,D., Galagan,J., Gardyna,S., Gilbert,D., Grant,G.,
            Hagos,B., Heaford,A., Horton,L., Howland,J.C., Jones,C., Kann,L.,
            Karatas,A., Lehoczky,J., Lieu,C., Locke,K., Macdonald,P.,
            Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J.,
            Meldrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
            Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B.,
            Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P.,
            Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A.,
            Tesfaye,S., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A.,
            Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-AUG-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Apr 8, 2000 this sequence version replaced gi:6751797.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L2077
                Center clone name: 173_J_18
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-primer-amersham; 4% of reads
                Chemistry: Dye-terminator Big Dye; 96% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 154528 bases at least Q40
                Consensus quality: 164381 bases at least Q30
                Consensus quality: 169670 bases at least Q20
                Insert size: 180000; agarose-fp
                Insert size: 175815; sum-of-contigs
                Quality coverage: 3.9 in Q20 bases; agarose-fp
                Quality coverage: 3.9 in Q20 bas.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 28 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1206: contig of 1206 bp in length
            *     1207 1306: gap of      100 bp
            *     1307     2316: contig of 1010 bp in length
            *     2317 2416: gap of      100 bp
            *     2417     3791: contig of 1375 bp in length
            *     3792 3891: gap of      100 bp
            *     3892     5035: contig of 1144 bp in length
            *     5036 5135: gap of      100 bp
            *     5136     6291: contig of 1156 bp in length
            *     6292 6391: gap of      100 bp
            *     6392     7820: contig of 1429 bp in length
            *     7821 7920: gap of      100 bp
            *     7921     9966: contig of 2046 bp in length
            *     9967 10066: gap of      100 bp
            *    10067    11414: contig of 1348 bp in length
            *    11415 11514: gap of      100 bp
            *    11515    13580: contig of 2066 bp in length
            *    13581 13680: gap of      100 bp
            *    13681    15919: contig of 2239 bp in length
            *    15920 16019: gap of      100 bp
            *    16020    18054: contig of 2035 bp in length
            *    18055 18154: gap of      100 bp
            *    18155    20599: contig of 2445 bp in length
            *    20600 20699: gap of      100 bp
            *    20700    25017: contig of 4318 bp in length
            *    25018 25117: gap of      100 bp
            *    25118    29998: contig of 4881 bp in length
            *    29999 30098: gap of      100 bp
            *    30099    34538: contig of 4440 bp in length
            *    34539 34638: gap of      100 bp
            *    34639    39508: contig of 4870 bp in length
            *    39509 39608: gap of      100 bp
            *    39609    46103: contig of 6495 bp in length
            *    46104 46203: gap of      100 bp
            *    46204    52117: contig of 5914 bp in length
            *    52118 52217: gap of      100 bp
            *    52218    58555: contig of 6338 bp in length
            *    58556 58655: gap of      100 bp
            *    58656    65212: contig of 6557 bp in length
            *    65213 65312: gap of      100 bp
            *    65313    72795: contig of 7483 bp in length
            *    72796 72895: gap of      100 bp
            *    72896    78614: contig of 5719 bp in length
            *    78615 78714: gap of      100 bp
            *    78715    89051: contig of 10337 bp in length
            *    89052 89151: gap of      100 bp
            *    89152   100736: contig of 11585 bp in length
            *   100737 100836: gap of      100 bp
            *   100837   117851: contig of 17015 bp in length
            *   117852 117951: gap of      100 bp
            *   117952   135700: contig of 17749 bp in length
            *   135701 135800: gap of      100 bp
            *   135801   153915: contig of 18115 bp in length
            *   153916 154015: gap of      100 bp
            *   154016   178515: contig of 24500 bp in length.
FEATURES             Location/Qualifiers
     source          1. .178515
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="8"
                     /map="8"
                     /clone="RP11-173J18"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .1206
                     /note="assembly_fragment"
     misc_feature    1307. .2316
                     /note="assembly_fragment"
     misc_feature    2417. .3791
                     /note="assembly_fragment"
     misc_feature    3892. .5035
                     /note="assembly_fragment"
     misc_feature    5136. .6291
                     /note="assembly_fragment"
     misc_feature    6392. .7820
                     /note="assembly_fragment"
     misc_feature    7921. .9966
                     /note="assembly_fragment"
     misc_feature    10067. .11414
                     /note="assembly_fragment"
     misc_feature    11515. .13580
                     /note="assembly_fragment"
     misc_feature    13681. .15919
                     /note="assembly_fragment"
     misc_feature    16020. .18054
                     /note="assembly_fragment"
     misc_feature    18155. .20599
                     /note="assembly_fragment"
     misc_feature    20700. .25017
```

```
                    /note="assembly_fragment"
     misc_feature   25118..29998
                    /note="assembly_fragment"
     misc_feature   30099..34538
                    /note="assembly_fragment"
     misc_feature   34639..39508
                    /note="assembly_fragment"
     misc_feature   39609..46103
                    /note="assembly_fragment"
     misc_feature   46204..52117
                    /note="assembly_fragment"
     misc_feature   52218..58555
                    /note="assembly_fragment"
     misc_feature   58656..65212
                    /note="assembly_fragment
                    clone_end:SP6
                    vector_side:right"
     misc_feature   65313..72795
                    /note="assembly_fragment"
     misc_feature   72896..78614
                    /note="assembly_fragment"
     misc_feature   78715..89051
                    /note="assembly_fragment"
     misc_feature   89152..100736
                    /note="assembly_fragment
                    clone_end:T7
                    vector_side:right"
     misc_feature   100837..117851
                    /note="assembly_fragment"
     misc_feature   117952..135700
                    /note="assembly_fragment"
     misc_feature   135801..153915
                    /note="assembly_fragment"
     misc_feature   154016..178515
                    /note="assembly_fragment"
BASE COUNT    54258 a  34971 c  35239 g  51342 t   2705 others
ORIGIN

  Query Match             41.0%;  Score 2328;  DB 74;  Length 178515;
  Best Local Similarity   71.9%;  Pred. No. 0;
  Matches 3790;  Conservative    0;  Mismatches 1140;  Indels  338;  Gaps   50;

Qy          1 CCCGGGGCACCACTGGCATCCCTCCCCCTACCACCCCCAATCCCTCCCTTTACGCCACCC 60
              ||  |  | ||| |    | |||     ||     |    |  |  ||||| |||     |  |  ||
Db      73670 CCATGAGAACCCCCATCCTCCCCAGGCCCCACTACTCTCATCCCACCCCACCCCCACCCT 73729

Qy         61 ATCCAAACATCTTCACGCTCACCCCCAGCCCAAGCCAGGCAGAATCCGGTTCCACCCCTG 120
               | | |     | | | | |         |  | ||||||||||| | ||||||
Db      73730 CATCTACCCTACCCTCACCCCCCTATCCCTACATTCTGGTAGAATCCGGTTTCGCCCCTG 73789

Qy        121 CTCTCAACCCAGGGAAG-CCCAGGTGCCCAGATGTGACGCCACTGACTTGAGCATTAGTG 179
              || |||||| ||||||  ||| ||||||| |||||||| |||||||||| || || |  |
Db      73790 CTTTCAACCGAGGGAATCCCCGGGTGCCCGGATGTGACACCACTGACTTG-GCGTTGGGG 73848

Qy        180 GTTAGAGAGAAGCGAGGTTTTCGGTCTGAGGGGCGGCTTGAGATCGGTGGAGGGAAGCGG 239
              || | ||||||||| ||| ||| |||||| |||||||||||| ||||||||||| ||
Db      73849 GTCAAAGAGAAGCGAAGTTCTCGCTCTGAGTGGCGGCTTGAGATTGGTGGAGGGAAGTGG 73908

Qy        240 GCCCA-GCTCTGTAAGGAGGCAAGGTGACATGCTGAGGGAGGACTGAGGACCCACTTACC 298
              | ||| |||||| | |||||||||||| |  || ||||||||| ||||  |  | |||
Db      73909 GTCCAGGCTCTGTGAAGAGGCAAGGTGAGACTCTAAGGGAGGACTCAGGAGGCCCCCACC 73968

Qy        299 CCAGATAGAGGACCCAAATAATCC---------CTTCATGCCAGTCCTGGACCATCTGG 349
              ||||||||||| |||||||||||          |  | |||||| |||  |||| ||||
Db      73969 CCAGATAGAGGGCCCCAAATAATCCAGCACTACTCCTGCTGCCAGCCCTAAACCACCTGG 74028

Qy        350 TGGTGGACTTCTCAGGCTGGGCCACCCCCAGCCCCCT---------TGCTGCTTAAACCA 400
               || ||||||||||| |||||||| ||| || ||||          ||  |||||| ||
Db      74029 GGGCGGACTTCTCAGTCTGGGCCACTCCCTGCACCCTGCCAGCCCCTGTGGCTTAAGCCG 74088

Qy        401 CTGGGGACTC-GAAGTCAGAGCTCCGTGTGATCAGGGAAGGGCTGCTTAGGAGAGGGCAG 459
              |  ||||||| | |||||||||| |||||| || || ||||||| |||||||||||||||
Db      74089 CAAGGGACTCTGGAGTCAGAGCTTCGTGTGACCAAGGCAGGGCTGGTTAGGAGAGGGCAG 74148

Qy        460 CGTCCAGGCTCTGCCAGACATCATGCTCAGGATTCTCAAGGAGGGCTGAGGGTCCCTAAG 519
               | ||| |||||||||| |||||    |||||| || |  ||||||||||| ||| ||
Db      74149 TGGCCAAGCTCTGCCAGGCATCAACATCAGGACCCTAAGTGAGGGCTGAGGGCCCCCAAC 74208

Qy        520 ACCC---CACTCCCGTGACCCAACCCCCACTCCAATGCTCACTCCCGTGACCCAACCCCC 576
               |||   | |  ||   |||| | ||||| |  | || |  | ||  |||    ||  ||||
Db      74209 CCCCATTCCCATCCCCCACCCCATCCCCATTTCCATCCCCATCCCCACCTCCATCCCCCA 74268

Qy        577 TCTTCATTGTCATTCCAACCCCCACCCCACATCCCCCACCCCATCCCTCAACCCTGATGC 636
               |   |    || ||   ||  |   ||   ||||  | |  |    ||  ||        |
Db      74269 CCAGAACCCCTATCCCCCACCAGAAACCCTATCCTGCCCACTCCCCACCACGTTCATCCC 74328

Qy        637 CCATCCGCCCAGCCATTCCACCCTCACCCC-CACCCCCACCCCCACGCCCACTCCCACCC 695
               |||  |||   |      |  || ||||||    ||||  |  || |||  |  | |||  |||
Db      74329 TACTCCCACCATTCCCATCTTCCCCACCCCTGACCCTCCTCCTCACCCTCCCACCCCCCC 74388

Qy        696 CCACCCAGGCAGGATCCGG------------TTCCCGCCAGGAAACATCCGGGTGCCCG 742
               ||||| ||||| || |||            |||   ||||| ||   | ||||| ||
Db      74389 ACACCCTGGCAGAATTCGGTTCTGCTCCTGCTTTCAACCCAGGGAATCCCTGGGTGACCA 74448

Qy        743 GATGTGACGCCACTGACTTGCGCATTGTGGGGCAGAGAGAAGCGAGGTTTCCATTCTGAG 802
              |||||| ||||||| ||||| ||||      || | ||||||||| |||     |  ||| ||
Db      74449 GATGTGGTGCCACTGTCTTGCACATTTGAGGTCGGAGAGAAGCAAGGGCCTCGCTCTCAG 74508

Qy        803 GGACGGCGTAGAGTTCGGCCGAAGGAACCTGACCCAGGCTCTGTGAGGAGGCAAGGTGAG 862
              || | || | ||| || || | || || ||  |  ||  ||||||||||||| ||||||||
Db      74509 GGGCAGC-TGGAGATCAGCTGAGGGCAGCTGGCCCTGGCTCTGTGAGGATGCAAGGTGAG 74567

Qy        863 AGGCTGAGGGAGGACTGAGGACCCCGCCACTCCAAATAGAGAGCCCCAAATATTCCAGCC 922
              | ||||||||||||| ||||     |||| ||    ||| |  ||||||||| |||||
Db      74568 ATGCTGAGGGAGGACTAAGGAGTATCCCACCCCTGGTAGTGGACCCCAAATAATCCAGTG 74627

Qy        923 CCGCC---CTTGCTGCCAGCCCTGGCCCACCCGCGGGAAGACGTCTCAGCCTGGGCTGCC 979
              || ||    | |||||| ||| |||| ||| ||   || | |||   || || |||||| ||
Db      74628 CCACCTCTCCTGCTGCTAGCTCTGGACCATCCAGGGCAGGACTCCTTAGGCTGGGCCACC 74687

Qy        980 CCCAGACCCCTGCTCCAAAAGCCTTGAGAGACACCAGGTTCTTCTCCCCAAGCTCTGGAA 1039
              ||||| ||||  |    | | ||||                    ||| |||
Db      74688 CCCAGTCCCCCACCGCTTAAGCCGCAGGGGA--------------------CTCAGGAG 74726

Qy       1040 TCAGAGGTTGCTGTGACCAGGGCAGGACTGGTTAGGAGAGGGCAGGGCACAGGCTCTGCC 1099
               ||||| |||  | |||||||||||||||||||||||||||| |||   | ||||||||||
Db      74727 ACAGAGCTTGGTATGACCAGGGCAGGACTGGTTAGGAGAGGACAGCTCCCAGGCTCTGCC 74786

Qy       1100 AGGCATCAAGATCAGCACCCAAGAGGGAGGGCTGTGGGCCCCCAAGACTGCACTCCAATC 1159
              ||| | |||  ||||  | || |   || | |       |     ||||
Db      74787 AGGAAACAACGTCAGGAACCTAAGGGAAAGCTGAGGCTACCCC----------------- 74829

Qy       1160 CCCACTCCCACCCCATTCGCATTCCCATTCCCCACCCAACCCCCATCTCCTCAGCTACAC 1219
                                                                         |||
Db      74830 ---------------------------------------------------------CAC 74832

Qy       1220 CTCCACCCCCATCCCTACTCCTACTCCGTCACCTGACCACCACCCTCCAGCCCCAGCACC 1279
               | | | || |||   |||||  |      ||  |||  ||||| |||  ||      ||
Db      74833 GCCAAACTCTATTCCTGTCCCTACCTCCGTCCCCCACCTACACCCCCATTCC-----CC 74887

Qy       1280 AGCCCCAACCCTTCTGCCACCTCACCCTCACTGCCCCCAACCCCACCCTCATCTCTCTCA 1339
               ||||| ||||  |  | ||||||    | |||  ||||||
Db      74888 CACCCCATCCCTACCGGCACCTCTATCCCACATCCCCCA-------------------- 74926

Qy       1340 TGTGCCCCACTCCCATCGCCTCCCCCATTCTGGCAGAATCCGGTT-TGCCCCTGCTCTCA 1398
                                 |||| ||  ||||||||||||| || |||||||| |  ||
Db      74927 --------------------CCCCTATCCTGGCAGAATCCGATTCTGCCCCTGATTTCA 74965

Qy       1399 ACCCAGGGAAGCCCTGGTAGGCCCGATGTGAAACCACTGACTTGAACCTCACAGATCTGA 1458
              ||||||||||||||  | ||||  |||||  |    |||||||||  |  |  | | | | ||
Db      74966 ACCCAGGGAAGCCCTAGGGGGCCGGATGTGATGCTGCTGACTTGTGCATTGCGGGTCAGA 75025

Qy       1459 GAGAAGCCAGGTTCATTTAATGGTTCTGAGGGGCGGCTTGAGATCCACTGAGGGGAGTGG 1518
              ||||| | |||           ||||||||||| || |  ||||||||  | ||||||| |||
Db      75026 GAGAATCAAGG------GCATGGTTCTGAGAAGCCGACTGAGATCAGCAGAGGGGAATGG 75079
```


Qy 1519 TTTTAGGCTCTGTGAGGAGGCAAGGTGAGA-TGCTGAGGGAGGACTGAGGAGGCACACAC 1577
Db 75080 GCCCGGGCTCTGTGAGGAGGCAAGGTGAGACCCCCGAGGAAGGAATGAGGAAGCCCTCAC 75139
Qy 1578 CCCAGGTAGATGGCCCCAAAATGATCCAGTACCACCCCTGCTGCCAGCCCTGGACCACCC 1637
Db 75140 CC--AGATAGAGAACCCCAAATAATCCAGTACTACCTTTGCTGCCAGCCCTGGACCAC-- 75195
Qy 1638 GGCCAGGACAGATGTCTCAGCTGGACCACCCCCCGTCCCGTCCCACTGCCACTTAACCCA 1697
Db 75196 --CCAGGGCAGACTTCTCAGGCTGGACCTTCCC---CCCTCCCCACTGCCACTTAAGCCA 75250
Qy 1698 CAGGGCAATCTGTAGTCATAGCTT-ATGTGACCGGGGCAGGGTTGGTCAGGAGAGGCAGG 1756
Db 75251 CAAGGGACTCTGGAGTCAGACCTTGGTGTGACCAGGGAAGGGCCGGTCAGGAGAGG---- 75306
Qy 1757 GCCCAGGCATCAAGGTCCAGCATCCGCCCGGCATTAGGGTCAGGACCCTGGGAGGGAACT 1816
Db 75307 ---------GCAGGGGCCAGGCTCTGTCAGGCATCAAAATCAGGACCCTGAGAGAGAATT 75357
Qy 1817 GAGGGTTCCCCACCCACACCTGTCTCCTCATCTCCACCGCCACCCCACTCACATTCCCAT 1876
Db 75358 GAGGGCCCCCACCCCAACCCCTATACCCATCCCTAACCCCATACCCACTCTACTTGCATT 75417
Qy 1877 ACCTACC----CCCTACCCCCAACCTCATCTTGTCAGAA----------TCCCTGCTGTC 1922
Db 75418 CCCAGCCCCATCCCCACACCCTACCCCATCTTGGCAGAATCTGTTTCTTTCCCTGCAGTC 75477
Qy 1923 AACCCACGGAAGCCACGGGAATGGCGGCCAGGCACTCGGATCTTGACGTCCCCATCCAGG 1982
Db 75478 AACCCACAGAAGCCCCAGGAATGACAGACAGGCACACCCATTCTGACGTCCACATCCAGG 75537
Qy 1983 GTCTGATGGAGGGAAGGG-------GCTTGAACAGGGCCTCAGGGGAGCAGAGG----GA 2031
Db 75538 GCTGAAGGAGGGAAAGGGCTTAGTATCATGAGCAGGGCCTCAGGGGAGTCTCTGCTCCTC 75597
Qy 2032 GGGCCCTACTGCGAGATGAGGGAGGCCTCAGAGGACCCAGCACCCTAGGA---------C 2082
Db 75598 AAGCCCTGCTGGGAGTAAAGGGAGGCCTCAGGGAACCCAGGTCCTCAGGATAGGGGGTCC 75657
Qy 2083 ACCGCACCCCTGTCTGAGACTGAG--GCTGCCACTTCTGGCCTCAAGAATCAGAACGATG 2140
Db 75658 ACTCCAACCCTGTCTGAGACTGAGGCGCCTCCTCTTTCATCCTCGGGAATCAGAGGGATG 75717
Qy 2141 GGGACTCAGATTGCATGGGGGTGGGACCCAGGCCTGCAAGGCTTACGCGGAGGAAGAGGA 2200
Db 75718 GAGACTCACGTCAGCAGAGGGTGGGGCCCAACCCTGCCAGGATCAAGGAGAGGAAGAAGA 75777
Qy 2201 GGGAGGACTCAGGGGACCTTGGAATCCAGATCAGTGTGGACCTCGGCCCTGAGAGGTCCA 2260
Db 75778 GGGAGGACTCAGGGTACCTTTGAGTCCAGAACAATGGGGACCTTTGCCCTGGGAGGTCCA 75837
Qy 2261 GGGCACGGTGGCCACATATGGCCCATATTTCCTGCATCTTTGAGGTGAC----AGGACAG 2316
Db 75838 GTGCACAGTGGCCACCTGTAGCCCATGCTTGCTGCACCTTCTGGGTGACAAAGAGGAGAG 75897
Qy 2317 AGCTGTGGTCTGAGAAGTGGGGCCTCAGGTCAACAGAGGGAGGAGTTCCAGGATCCATAT 2376
Db 75898 GGCTGTGGTCAGAGCAGTGGTGACTCAGGTCAGCAGAGGGAGGAGTCCCAGCATCTGCAG 75957
Qy 2377 GGCCCAAGATGTGCCCCCTTCATGAGGACTGGGGATATCCCCGGCTCAGAAAGAAGGGAC 2436
Db 75958 GCCCCAATGTGTGCCCCATTCATGAAGATTGGGGATA-CCTTGGCTCAGAAAGAAGGGAC 76016
Qy 2437 TCCACACAGTCTGGCTGTCCCCTTTTAGTAGCTCTAGGGGGACCAGATCAGGGATGGCGG 2496
Db 76017 CCCACAGAGTCTGGCTGTCCCCTGATTTTTGCTCAGAGGGGACCAAATCAAGGATAGCCC 76076
Qy 2497 TATGTTCCATTCTCACTTGTACCACAGGCAGGAAGTTGGGGGGCCCTCAGGGAGATGGGG 2556
Db 76077 TATGTGCCAACCTCATTTGTGCCACAGGAAAGAAGTTGAAGAGCCCTCAGGGTGATGGGG 76136
Qy 2557 TCTTGGGGTAAAGGGGGGATGTCTACTCATGTCAGGGAATTGGGGGTTGAGGAAGCACAG 2616
Db 76137 TCTTGCAGTAAAGGGGAGCTATCTGCTCATCTCAGGGGGTTTCAGGTTGAGGAATGGCAG 76196
Qy 2617 GCGCTGGCAGGAATAAAGATGAGTGAGACAGACAAGGCTATTGGAATCCACACCCCAGAA 2676
Db 76197 GCCCCATCACGATGAA----GAGTAACCCACAGGAGGCCATAGAAACACTCACCCCAGAA 76252
Qy 2677 CCAAAGGGGTCAGCCCTGGACACCTCA---------CCCAGGATGTGGCTTCTTTTTCAC 2727
Db 76253 CCAAAGGGGTCATACCTGGACACCCCATGTGGGGGTGACAGGATGTAGC-TCCATCTCAT 76311
Qy 2728 TCCTGTTTCCAGATCTGGGGCAGGTGAGGACCTCATTCTCAGAGGGTGACTCAGGTCAAC 2787
Db 76312 TCCTGTTTTCAGATCTCGGGGAGGTGAGGAACTTGTTCTCCGAGGATGACTCAGGTCAAC 76371
Qy 2788 GTAGGGACCCCCATCTGGTCTAAAGACAGAGCGGTCCCAGGATCTGCCATGCGTTCGGGT 2847
Db 76372 ACAGGGGCCCCCATCTGGTGGATAGACAGAGTGGTCCCAGGATCTGTCAGTAGTTCCGGT 76431
Qy 2848 GAGGAACATGAGGGAGGACTGAGGGTACCCCAGGACCAGAACACTGAGGGAGACTGCACA 2907
Db 76432 GAGGAACATGAGGGACGATTGAGGGCACCCTTGGGCCAGAACACAGATGAGGACCTCACG 76491
Qy 2908 GAAATCAGCCCTGCCCCTGCTGTCACCCCAGAGAGCATGGGCTGGGCCGTCTGCCGAGGT 2967
Db 76492 GAAATCTGCCCTGCCCCTGCTGTCACTCCAGAGAGCATGGGCAGGGCTGTCTGCTGCAGT 76551
Qy 2968 CCTTCCG--TTATCCTGGGATCATTGATGTCAGGGACGGGGAGGCCTTGGTCTGAGAAGG 3025
Db 76552 CCCCCCGACTTACCCTGGGATCATTGGTGTCAGGGATGGGGAGGTCTTTGTC-GAGGGGT 76610
Qy 3026 CTGCGCTCAGGTCAGTAGAGGGAGCGTCCCAGGCCCTGCCAGGAGTCAAGGTGAGGACCA 3085
Db 76611 CTGCACTCAGGTCAGTAGAGGGAGCGTCTTAGGCCCTGCCAGGAGACAAGGTAAGAACGA 76670
Qy 3086 AGCGGGCACCTCACCCAGGACACATTAATTCCAATGAATTTTGATATCTCTTGCTGCCCT 3145
Db 76671 AGCAGGTTCCTCACCCAGGACACATGAATTCCAATGCATTTCAGCATCTCTTCCTGTCCT 76730
Qy 3146 TCCCCA-AGGACCTAGGCACGTGTGGCCAGATGTTTGTCCCCTCCTGTCCTTCCATTCCT 3204
Db 76731 TCCCAAGAGGACCTGGGCACGTGTGGCCAGATGTGAGTCTCCTCATGTCCT---GTTCCC 76787
Qy 3205 TATCATGGATGTGAACTCTTGATTTGGATTTCTCAGACCAGCAAAAGGGCAGGATCCAGG 3264
Db 76788 TATCAGGGATGTGAGCTCTTAATCTGAGTTTCTCAGGCCAGCAAAAGGGTGGGATCCAGG 76847
Qy 3265 CCCTGCCAGGAAAAATATAAGGGCCCTGCGTGAGAACAGAGGGGGTCATCCACTGCATGA 3324
Db 76848 CCTTGCCAGGAGAAAGGTGAGGGCCCTGTGTGAGCACAGAGGGGACCATTCACCCCAAGA 76907
Qy 3325 GAGTGGGGATGTCACAGAGTCCAGCCCACCCTCCTGGTAGCACTGAGAAGCCAGGGCTGT 3384
Db 76908 GGGTGGAGACCTCACAGATTCCAGCCTACCCTCCTGTTAGCACTGGGGGCCTGAGGCTGT 76967
Qy 3385 GCTTGCGGTCTGCACCCTGAGGGCCCGTGGATTCCTCTTCCTGGAGCTCCAGGAACCAGG 3444
Db 76968 GCTTGCAGTCTGCACCCTGAGGGCCCATGCATTCCTCTTCCAGGAGCTCCAGGAAACAGA 77027
Qy 3445 CAGTGAGGCCTTGGTCTGAGACAGTATCCTCAGGTCACAGAGCAGAGGATGCACAG---- 3500
Db 77028 CACTGAGGCCTTGGTCTGAGGCCGTGCCCTCAGGTCACAGAGCAGAGGAGATGCAGACGT 77087
Qy 3501 GGTGTGCCAGCAGTGAATGTTTGCCCTGAATGCACACCAAGGGCCCCACCTGCCACAGGA 3560
Db 77088 CTAGTGCCAGCAGTGAACGTTTGCCTTGAATGCACACTAATGGCCCCCATCGCCCCAGAA 77147
Qy 3561 CACATAGGACTCCACAGAGTCTGGCCTCA-CCTCCCTACTGTCAGTCCTGTAGAATCGAC 3619
Db 77148 CATATGGGACTCCAGAGCACCTGGCCTCACCCTCTCTACTGTCAGTCCTGCAGAATCAGC 77207
Qy 3620 CTCTGCTGGCCGGCTGTACCCTGAGTACCCTCTCACTTCCTCCTTCAGGTTTTCAGGGGA 3679



Search completed: September 13, 2000, 05:33:47
Job time: 13929 sec

****

****

Release 3.1A John F. Collins, Biocomputing Research Unit.

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Sep 12 13:19:06 2000; MasPar time 5.54 Seconds
76.608 Million cell updates/sec

Tabular output not generated.

Title: >US-08-819-669E-26
Description: (1-9) from US08819669E.pep
Perfect Score: 61
Sequence: 1 EADPTGHSY 9

Scoring table: PAM 150
Gap 15

Searched: 142080 seqs, 47172406 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: pir64
1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 20.662; Variance 21.178; scale 0.976

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	61	100.0	280	2	JC2358	tumor-associated anti	1.51e-04
2	54	88.5	234	2	I38667	MAGE-8 antigen - huma	1.43e-02
3	54	88.5	315	2	I38668	MAGE-9 antigen - huma	1.43e-02
4	52	85.2	369	2	I38659	MAGE-10 antigen - hum	5.00e-02
5	51	83.6	319	2	I38660	MAGE-11 antigen - hum	9.24e-02
6	48	78.7	129	2	E72685	hypothetical protein	5.60e-01
7	48	78.7	269	2	A49334	Ras homolog Rad - hum	5.60e-01
8	48	78.7	925	1	A39216	plasma cell membrane	5.60e-01
9	47	77.0	497	1	S33938	penton protein (III)	1.01e+00
10	46	75.4	9	2	PH1299	MAGE 5 protein - huma	1.79e+00
11	46	75.4	314	2	JC2360	tumor-associated anti	1.79e+00
12	45	73.8	1033	2	S02168	type I site-specific	3.17e+00
13	44	72.1	98	2	F70769	hypothetical protein	5.55e+00
14	44	72.1	128	2	B72600	hypothetical protein	5.55e+00
15	44	72.1	461	2	T09933	nucleotide pyrophosph	5.55e+00
16	44	72.1	488	1	S55874	sulfite oxidase (EC 1	5.55e+00
17	44	72.1	488	1	A53107	sulfite oxidase (EC 1	5.55e+00
18	44	72.1	725	2	A45033	myelin transcription	5.55e+00
19	43	70.5	197	2	A70832	hypothetical protein	9.63e+00
20	43	70.5	290	2	S64312	hypothetical protein	9.63e+00
21	43	70.5	314	2	JC2361	tumor-associated anti	9.63e+00
22	43	70.5	347	2	I38008	MAGE-Xp protein - hum	9.63e+00
23	43	70.5	669	2	I38029	matrix metalloprotein	9.63e+00
24	43	70.5	745	2	T10924	3C3.15c protein - Str	9.63e+00
25	43	70.5	875	2	T12794	hypothetical yomG pro	9.63e+00
26	43	70.5	878	2	S44543	hypothetical protein	9.63e+00
27	42	68.9	156	2	B36905	conserved hypothetica	1.66e+01
28	42	68.9	283	2	E69626	methylenetetrahydrofo	1.66e+01
29	42	68.9	288	2	A56279	carbon-monoxide dehyd	1.66e+01
30	42	68.9	301	2	C71194	hypothetical protein	1.66e+01
31	42	68.9	317	2	JC2359	tumor-associated anti	1.66e+01
32	42	68.9	326	2	S44753	C13G5.2 protein - Cae	1.66e+01
33	42	68.9	348	1	DEEBOT	dihydroorotase (EC 3.	1.66e+01
34	42	68.9	461	2	A70687	probable oxidoreducta	1.66e+01
35	42	68.9	503	2	H70509	hypothetical protein	1.66e+01
36	42	68.9	503	2	G70506	hypothetical protein	1.66e+01
37	42	68.9	508	1	ISRTSS	protein disulfide-iso	1.66e+01
38	42	68.9	555	1	RGASWA	regulatory protein we	1.66e+01
39	42	68.9	700	2	S38928	translation elongatio	1.66e+01
40	42	68.9	747	2	B47093	cellulase (EC 3.2.1.4	1.66e+01
41	42	68.9	1052	2	T14343	zinc finger RNA bindi	1.66e+01
42	42	68.9	3396	1	A42551	genome polyprotein -	1.66e+01
43	41	67.2	366	2	A53286	cell-surface glycopro	2.82e+01
44	41	67.2	637	2	T03842	fission yeast Skb1 pr	2.82e+01
45	41	67.2	3898	1	A44217	genome polyprotein -	2.82e+01

ALIGNMENTS

RESULT 1
ENTRY JC2358 #type complete
TITLE tumor-associated antigen MAGE-1 - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 04-Sep-1998
ACCESSIONS JC2358
REFERENCE JC2358
#authors Ding, M.; Beck, R.J.; Keller, C.J.; Fenton, R.G.
#journal Biochem. Biophys. Res. Commun. (1994) 202:549-555
#title Cloning and analysis of MAGE-1-related genes.
#cross-references MUID:94311935
#accession JC2358
##molecule_type mRNA
##residues 1-280 ##label DIN
##experimental_source melanoma cell line DM150
GENETICS
#gene MAGE
CLASSIFICATION #superfamily tumor associated protein MAGE
FEATURE
161-169 #region HLA-A1 binding #status predicted
SUMMARY #length 280 #molecular-weight 30932 #checksum 467

Query Match 100.0%; Score 61; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 1.51e-04;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Db    161 EADPTGHSY 169
          |||||||||
Qy      1 EADPTGHSY 9
```

RESULT 2
ENTRY I38667 #type complete
TITLE MAGE-8 antigen - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 24-Sep-1999
ACCESSIONS I38667
REFERENCE I38659
#authors De Plaen, E.; Arden, K.; Traversari, C.; Gaforio, J.J.; Szikora, J.P.; De Smet, C.; Brasseur, F.; van der Bruggen, P.; Lethe, B.; Lurquin, C.; Brasser, R.; Chomez, P.; De Backer, O.; Cavenee, W.; Boon, T.
#journal Immunogenetics (1994) 40:360-369
#title Structure, chromosomal localization, and expression of 12 genes of the MAGE family.

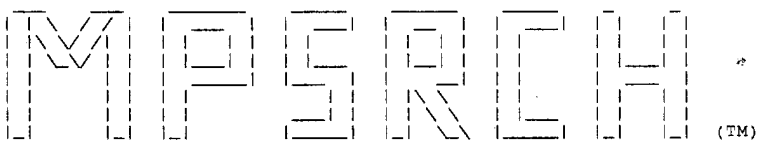
```
   #cross-references MUID:95012457
   #accession    I38667
      ##status        preliminary; translated from GB/EMBL/DDBJ
      ##molecule_type DNA
      ##residues      1-234 ##label RES
      ##cross-references EMBL:U10693; NID:g533525; PIDN:AAA68876.1;
                        PID:g533526
GENETICS
   #gene         GDB:MAGEA8; MAGE8
      ##cross-references GDB:331123
   #map_position Xq28-Xq28
   #introns      #status absent
CLASSIFICATION   #superfamily tumor associated protein MAGE
SUMMARY          #length 234  #molecular-weight 25197  #checksum 311

  Query Match           88.5%;  Score 54;  DB 2;  Length 234;
  Best Local Similarity  77.8%;  Pred. No. 1.43e-02;
  Matches     7;  Conservative    1;  Mismatches   1;  Indels   0;  Gaps   0;

Db    171 EVDPAGHSY 179
          | ||:||||
Qy      1 EADPTGHSY 9


RESULT    3
ENTRY            I38668     #type complete
TITLE            MAGE-9 antigen - human
ORGANISM         #formal_name Homo sapiens #common_name man
DATE             07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change
                   24-Sep-1999
ACCESSIONS       I38668
REFERENCE        I38659
   #authors      De Plaen, E.; Arden, K.; Traversari, C.; Gaforio, J.J.;
                   Szikora, J.P.; De Smet, C.; Brasseur, F.; van der Bruggen,
                   P.; Lethe, B.; Lurquin, C.; Brasser, R.; Chomez, P.; De
                   Backer, O.; Cavenee, W.; Boon, T.
   #journal      Immunogenetics (1994) 40:360-369
   #title        Structure, chromosomal localization, and expression of 12
                   genes of the MAGE family.
   #cross-references MUID:95012457
   #accession    I38668
      ##status        preliminary; translated from GB/EMBL/DDBJ
      ##molecule_type DNA
      ##residues      1-315 ##label RES
      ##cross-references EMBL:U10694; NID:g533527; PIDN:AAA68877.1;
                        PID:g533528
GENETICS
   #gene         GDB:MAGEA9; MAGE9
      ##cross-references GDB:331125
   #map_position Xp21.3-Xp21.3
   #introns      #status absent
CLASSIFICATION   #superfamily tumor associated protein MAGE
SUMMARY          #length 315  #molecular-weight 35088  #checksum 2468

  Query Match           88.5%;  Score 54;  DB 2;  Length 315;
  Best Local Similarity  77.8%;  Pred. No. 1.43e-02;
  Matches     7;  Conservative    1;  Mismatches   1;  Indels   0;  Gaps   0;

Db    167 EVDPAGHSY 175
          | ||:||||
Qy      1 EADPTGHSY 9


RESULT    4
ENTRY            I38659     #type complete
TITLE            MAGE-10 antigen - human
ORGANISM         #formal_name Homo sapiens #common_name man
DATE             07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change
                   24-Sep-1999
ACCESSIONS       I38659
REFERENCE        I38659
   #authors      De Plaen, E.; Arden, K.; Traversari, C.; Gaforio, J.J.;
                   Szikora, J.P.; De Smet, C.; Brasseur, F.; van der Bruggen,
                   P.; Lethe, B.; Lurquin, C.; Brasser, R.; Chomez, P.; De
                   Backer, O.; Cavenee, W.; Boon, T.
   #journal      Immunogenetics (1994) 40:360-369
   #title        Structure, chromosomal localization, and expression of 12
                   genes of the MAGE family.
   #cross-references MUID:95012457
   #accession    I38659
      ##status        preliminary; translated from GB/EMBL/DDBJ
      ##molecule_type DNA
      ##residues      1-369 ##label RES
      ##cross-references EMBL:U10685; NID:g533510; PIDN:AAA68869.1;
                        PID:g533511
GENETICS
   #gene         GDB:MAGEA10; MAGE10
      ##cross-references GDB:331126
   #map_position Xq28-Xq28
   #introns      #status absent
CLASSIFICATION   #superfamily tumor associated protein MAGE
SUMMARY          #length 369  #molecular-weight 40766  #checksum 586

  Query Match           85.2%;  Score 52;  DB 2;  Length 369;
  Best Local Similarity  77.8%;  Pred. No. 5.00e-02;
  Matches     7;  Conservative    1;  Mismatches   1;  Indels   0;  Gaps   0;

Db    193 EVDPTGHSF 201
          | ||||||:
Qy      1 EADPTGHSY 9


RESULT    5
ENTRY            I38660     #type complete
TITLE            MAGE-11 antigen - human
ORGANISM         #formal_name Homo sapiens #common_name man
DATE             07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change
                   24-Sep-1999
ACCESSIONS       I38660
REFERENCE        I38659
   #authors      De Plaen, E.; Arden, K.; Traversari, C.; Gaforio, J.J.;
                   Szikora, J.P.; De Smet, C.; Brasseur, F.; van der Bruggen,
                   P.; Lethe, B.; Lurquin, C.; Brasser, R.; Chomez, P.; De
                   Backer, O.; Cavenee, W.; Boon, T.
   #journal      Immunogenetics (1994) 40:360-369
   #title        Structure, chromosomal localization, and expression of 12
                   genes of the MAGE family.
   #cross-references MUID:95012457
   #accession    I38660
      ##status        preliminary; translated from GB/EMBL/DDBJ
      ##molecule_type DNA
      ##residues      1-319 ##label RES
      ##cross-references EMBL:U10686; NID:g533512; PIDN:AAA68870.1;
                        PID:g533513
GENETICS
   #gene         GDB:MAGEA11; MAGE11
      ##cross-references GDB:331128
   #map_position Xq28-Xq28
   #introns      #status absent
CLASSIFICATION   #superfamily tumor associated protein MAGE
SUMMARY          #length 319  #molecular-weight 35536  #checksum 9402

  Query Match           83.6%;  Score 51;  DB 2;  Length 319;
  Best Local Similarity  77.8%;  Pred. No. 9.24e-02;
  Matches     7;  Conservative    1;  Mismatches   1;  Indels   0;  Gaps   0;

Db    171 EVDPTSHSY 179
          | |||:|||
Qy      1 EADPTGHSY 9


RESULT    6
ENTRY            E72685     #type complete
TITLE            hypothetical protein APE0901 - Aeropyrum pernix (strain K1)
ORGANISM         #formal_name Aeropyrum pernix
DATE             20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change
```

```
                  20-Aug-1999
ACCESSIONS        E72685
REFERENCE         A72450
   #authors       Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.;
                    Haikawa, Y.; Jin-no, K.; Takahashi, M.; Sekine, M.; Baba,
                    S.; Ankai, A.; Kosugi, H.; Hosoyama, A.; Fukui, S.; Nagai,
                    Y.; Nishijima, K.; Nakazawa, H.; Takamiya, M.; Masuda, S.;
                    Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;
                    Kushida, N.; Oguchi, A.; Aoki, K.; Kubota, K.; Nakamura,
                    Y.; Nomura, N.; Sako, Y.; Kikuchi, H.
   #journal       DNA Res. (1999) 6:83-101
   #title         Complete genome sequence of an aerobic hyper-thermophilic
                    Crenarchaeon, Aeropyrum pernix K1.
   #cross-references MUID:99310339
   #accession     E72685
      ##status         preliminary
      ##molecule_type DNA
      ##residues       1-129 ##label KAW
      ##cross-references DDBJ:AP000060; NID:g5104188; PIDN:BAA79885.1;
                          PID:d1043671; PID:g5104570
      ##experimental_source strain K1
GENETICS
   #gene          APE0901
SUMMARY           #length 129  #molecular-weight 14303  #checksum 2150

  Query Match           78.7%;  Score 48;  DB 2;  Length 129;
  Best Local Similarity  85.7%;  Pred. No. 5.60e-01;
  Matches     6;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db    115 DPAGHSY 121
          ||:||||
Qy      3 DPTGHSY 9


RESULT    7
ENTRY             A49334     #type complete
TITLE             Ras homolog Rad - human
ALTERNATE_NAMES   Ras associated with diabetes (Rad)
ORGANISM          #formal_name Homo sapiens #common_name man
DATE              07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change
                    28-Aug-1998
ACCESSIONS        A49334
REFERENCE         A49334
   #authors       Reynet, C.; Kahn, C.R.
   #journal       Science (1993) 262:1441-1444
   #title         Rad: a member of the Ras family overexpressed in muscle of
                    type II diabetic humans.
   #cross-references MUID:94069319
   #accession     A49334
      ##status         preliminary
      ##molecule_type mRNA
      ##residues       1-269 ##label REY
      ##cross-references GB:L24564
KEYWORDS          alternative initiators; GTP binding; P-loop
FEATURE
   59-66               #region nucleotide-binding motif A (P-loop)\
   164-167             #region GTP-binding NKXD motif
SUMMARY           #length 269  #molecular-weight 29262  #checksum 9237

  Query Match           78.7%;  Score 48;  DB 2;  Length 269;
  Best Local Similarity  55.6%;  Pred. No. 5.60e-01;
  Matches     5;  Conservative    4;  Mismatches   0;  Indels   0;  Gaps   0;

Db     80 EAEAAGHTY 88
          ||:::||:|
Qy      1 EADPTGHSY 9


RESULT    8
ENTRY             A39216     #type complete
TITLE             plasma cell membrane glycoprotein PC-1 - human
CONTAINS          nucleotide pyrophosphatase (EC 3.6.1.9); phosphodiesterase I
                    (EC 3.1.4.1)
ORGANISM          #formal_name Homo sapiens #common_name man
DATE              10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change
                    10-Sep-1999
ACCESSIONS        A39216; S21706; S23587; S51030
REFERENCE         A39216
   #authors       Buckley, M.F.; Loveland, K.A.; McKinstry, W.J.; Garson, O.M.;
                    Goding, J.W.
   #journal       J. Biol. Chem. (1990) 265:17506-17511
   #title         Plasma cell membrane glycoprotein PC-1. cDNA cloning of the
                    human molecule, amino acid sequence, and chromosomal
                    location.
   #cross-references MUID:91009202
   #accession     A39216
      ##status         preliminary
      ##molecule_type mRNA
      ##residues       1-925 ##label BUC
      ##cross-references GB:J05654
REFERENCE         S21706
   #authors       Funakoshi, I.; Kato, H.; Horie, K.; Yano, T.; Hori, Y.;
                    Kobayashi, H.; Inoue, T.; Suzuki, H.; Fukui, S.; Tsukahara,
                    M.; Kajii, T.; Yamashina, I.
   #journal       Arch. Biochem. Biophys. (1992) 295:180-187
   #title         Molecular cloning of cDNAs for human fibroblast nucleotide
                    pyrophosphatase.
   #cross-references MUID:92246539
   #accession     S21706
      ##status         not compared with conceptual translation
      ##molecule_type mRNA
      ##residues       1-925 ##label FUN1
   #accession     S23587
      ##molecule_type protein
      ##residues       116-121;247-271,'X',273-275;279-280,'X',282-283;303-316;
                        362-364;449-465;482-525;529-534,'X',536-551,'X',553,
                        'X',555-556;597-606;'X',727-730;775-782;840-846,'XX',
                        849-852,'X',854-859 ##label FUN2
      ##note          it is uncertain whether Met-1 or Met-53 is the initiator
REFERENCE         S51030
   #authors       Belli, S.I.; Goding, J.W.
   #journal       Eur. J. Biochem. (1994) 226:433-443
   #title         Biochemical characterization of human PC-1, an enzyme
                    possessing alkaline phosphodiesterase I and nucleotide
                    pyrophosphatase activities.
   #cross-references MUID:95094801
   #accession     S51030
      ##status         preliminary
      ##molecule_type mRNA
      ##residues       1-80 ##label BEL
GENETICS
   #gene          GDB:PDNP1; M6S1; NPPS
      ##cross-references GDB:132615; OMIM:173335
   #map_position 6q22-6q23
CLASSIFICATION    #superfamily nucleotide pyrophosphatase; somatomedin B
                    homology
KEYWORDS          glycoprotein; phosphoprotein; phosphoric diester hydrolase;
                    transmembrane protein
FEATURE
   77-97               #domain transmembrane #status predicted #label TMM\
   104-144             #domain somatomedin B homology #label SBH1\
   145-188             #domain somatomedin B homology #label SBH2\
   179,285,341,477,
   578,585,643,700,
   731,748             #binding_site carbohydrate (Asn) (covalent) #status
                         predicted\
   256                 #binding_site AMP (Thr) (covalent) #status predicted
SUMMARY           #length 925  #molecular-weight 104924  #checksum 7446

  Query Match           78.7%;  Score 48;  DB 1;  Length 925;
  Best Local Similarity  66.7%;  Pred. No. 5.60e-01;
  Matches     6;  Conservative    3;  Mismatches   0;  Indels   0;  Gaps   0;

Db    374 EPDSSGHSY 382
          |:|::||||
Qy      1 EADPTGHSY 9
```

```
RESULT    9
ENTRY             S33938     #type complete
TITLE             penton protein (III) - human adenovirus 12
ORGANISM          #formal_name Mastadenovirus h12 #common_name human adenovirus
                    12
DATE              10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change
                    10-Sep-1999
ACCESSIONS        S33938
REFERENCE         S33928
   #authors       Sprengel, J.
   #submission    submitted to the EMBL Data Library, June 1993
   #accession     S33938
      ##status        preliminary
      ##molecule_type DNA
      ##residues      1-497 ##label SPR
      ##cross-references EMBL:X73487; NID:g313361; PIDN:CAA51887.1;
                        PID:g313372
CLASSIFICATION    #superfamily adenovirus penton protein
SUMMARY           #length 497  #molecular-weight 56393  #checksum 2182

  Query Match            77.0%;  Score 47;  DB 1;  Length 497;
  Best Local Similarity  66.7%;  Pred. No. 1.01e+00;
  Matches     6;  Conservative    2;  Mismatches   1;  Indels   0;  Gaps   0;

Db     310 ETDPKGRSY 318
           |:|| |:||
Qy       1 EADPTGHSY 9


RESULT   10
ENTRY             PH1299     #type fragment
TITLE             MAGE 5 protein - human (fragment)
ALTERNATE_NAMES   MAGE 51 protein
ORGANISM          #formal_name Homo sapiens #common_name man
DATE              30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change
                    03-Aug-1998
ACCESSIONS        PH1299; PH1300
REFERENCE         PH1294
   #authors       Traversari, C.; van der Bruggen, P.; Luescher, I.F.; Lurquin,
                    C.; Chomez, P.; Van Pel, A.; De Plaen, E.; Amar-Costesec,
                    A.; Boon, T.
   #journal       J. Exp. Med. (1992) 176:1453-1457
   #title         A nonapeptide encoded by human gene MAGE-1 is recognized on
                    HLA-A1 by cytolytic T lymphocytes directed against tumor
                    antigen MZ2-E.
   #cross-references MUID:93018875
   #accession     PH1299
      ##molecule_type DNA
      ##residues      1-9 ##label TRA
   #accession     PH1300
      ##molecule_type DNA
      ##residues      1-9 ##label TR2
SUMMARY           #length 9  #checksum 3660

  Query Match            75.4%;  Score 46;  DB 2;  Length 9;
  Best Local Similarity  66.7%;  Pred. No. 1.79e+00;
  Matches     6;  Conservative    3;  Mismatches   0;  Indels   0;  Gaps   0;

Db       1 EADPTSNTY 9
           |||||:::|
Qy       1 EADPTGHSY 9


RESULT   11
ENTRY             JC2360     #type complete
TITLE             tumor-associated antigen, MAGE 6 - human
ALTERNATE_NAMES   melanoma antigen 6; tumor-associated antigen, MAGE-3b
ORGANISM          #formal_name Homo sapiens #common_name man
DATE              20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change
                    24-Sep-1999
ACCESSIONS        JC2360; PH1301; I38665; G01445
REFERENCE         JC2358
   #authors       Ding, M.; Beck, R.J.; Keller, C.J.; Fenton, R.G.
   #journal       Biochem. Biophys. Res. Commun. (1994) 202:549-555
   #title         Cloning and analysis of MAGE-1-related genes.
   #cross-references MUID:94311935
   #accession     JC2360
      ##molecule_type mRNA
      ##residues      1-314 ##label DIN
      ##experimental_source melanoma cell line DM150
REFERENCE         PH1294
   #authors       Traversari, C.; van der Bruggen, P.; Luescher, I.F.; Lurquin,
                    C.; Chomez, P.; Van Pel, A.; De Plaen, E.; Amar-Costesec,
                    A.; Boon, T.
   #journal       J. Exp. Med. (1992) 176:1453-1457
   #title         A nonapeptide encoded by human gene MAGE-1 is recognized on
                    HLA-A1 by cytolytic T lymphocytes directed against tumor
                    antigen MZ2-E.
   #cross-references MUID:93018875
   #accession     PH1301
      ##molecule_type DNA
      ##residues      168-176 ##label TRA
REFERENCE         I38659
   #authors       De Plaen, E.; Arden, K.; Traversari, C.; Gaforio, J.J.;
                    Szikora, J.P.; De Smet, C.; Brasseur, F.; van der Bruggen,
                    P.; Lethe, B.; Lurquin, C.; Brasser, R.; Chomez, P.; De
                    Backer, O.; Cavenee, W.; Boon, T.
   #journal       Immunogenetics (1994) 40:360-369
   #title         Structure, chromosomal localization, and expression of 12
                    genes of the MAGE family.
   #cross-references MUID:95012457
   #accession     I38665
      ##status        preliminary; translated from GB/EMBL/DDBJ
      ##molecule_type DNA
      ##residues      1-314 ##label RES
      ##cross-references EMBL:U10691; NID:g533522; PIDN:AAA68875.1;
                        PID:g533523
REFERENCE         G07126
   #authors       Fenton, R.G.
   #submission    submitted to the EMBL Data Library, June 1994
   #accession     G01445
      ##status        preliminary; translated from GB/EMBL/DDBJ
      ##molecule_type mRNA
      ##residues      1-314 ##label FEN
      ##cross-references EMBL:U10339; NID:g499121; PIDN:AAA19006.1;
                        PID:g499122
GENETICS
   #gene          GDB:MAGEA6; MAGE6
      ##cross-references GDB:331121
   #map_position Xq28-Xq28
   #introns       #status absent
CLASSIFICATION    #superfamily tumor associated protein MAGE
FEATURE
   168-176            #region HLA-A1 binding #status predicted
SUMMARY           #length 314  #molecular-weight 34891  #checksum 9870

  Query Match            75.4%;  Score 46;  DB 2;  Length 314;
  Best Local Similarity  66.7%;  Pred. No. 1.79e+00;
  Matches     6;  Conservative    0;  Mismatches   3;  Indels   0;  Gaps   0;

Db     168 EVDPIGHVY 176
           | || || |
Qy       1 EADPTGHSY 9


RESULT   12
ENTRY             S02168     #type complete
TITLE             type I site-specific deoxyribonuclease (EC 3.1.21.3)
                    EcoR124/3 chain hsdR - Escherichia coli plasmid R124/3
ALTERNATE_NAMES   type I restriction enzyme EcoR124/3 chain hsdR
ORGANISM          #formal_name Escherichia coli
DATE              01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change
                    09-Sep-1997
ACCESSIONS        S02168
```

```
REFERENCE       S02166
   #authors     Price, C.; Lingner, J.; Bickle, T.A.
   #journal     J. Mol. Biol. (1989) 205:115-125
   #title       Basis for changes in DNA recognition by the EcoR124 and
                  EcoR124/3 type I DNA restriction and modification enzymes.
   #cross-references MUID:89178628
   #accession   S02168
      ##molecule_type DNA
      ##residues      1-1033 ##label PRI
      ##cross-references EMBL:X13145; NID:g388978; PID:g41750
GENETICS
   #gene        hsdR
   #genome      plasmid
KEYWORDS        DNA binding; hydrolase
SUMMARY         #length 1033  #molecular-weight 119656  #checksum 4436

  Query Match           73.8%;  Score 45;  DB 2;  Length 1033;
  Best Local Similarity  75.0%;  Pred. No. 3.17e+00;
  Matches     6;  Conservative    1;  Mismatches   1;  Indels   0;  Gaps   0;

Db      23 AEPTGDSY 30
           |:||| ||
Qy       2 ADPTGHSY 9


RESULT   13
ENTRY           F70769     #type complete
TITLE           hypothetical protein Rv1322 - Mycobacterium tuberculosis
                  (strain H37RV)
ORGANISM        #formal_name Mycobacterium tuberculosis
DATE            17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change
                  17-Jul-1998
ACCESSIONS      F70769
REFERENCE       A70500
   #authors     Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher,
                  C.; Harris, D.; Gordon, S.V.; Eiglmeier, K.; Gas, S.; Barry
                  III, C.E.; Tekaia, F.; Badcock, K.; Basham, D.; Brown, D.;
                  Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.;
                  Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
                  Hornsby, T.; Jagels, K.; Krogh, A.; McLean, J.; Moule, S.;
                  Murphy, L.; Oliver, S.; Osborne, J.; Quail, M.A.;
                  Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.;
                  Skelton, S.; Squares, S.; Sqares, R.; Sulston, J.E.;
                  Taylor, K.; Whitehead, S.; Barrell, B.G.
   #journal     Nature (1998) 393:537-544
   #title       Deciphering the biology of Mycobacterium tuberculosis from
                  the complete genome sequence.
   #cross-references MUID:98295987
   #accession   F70769
      ##status        preliminary; nucleic acid sequence not shown;
                       translation not shown
      ##molecule_type DNA
      ##residues      1-98 ##label COL
      ##cross-references GB:Z73902; GB:AL123456; NID:g3261576; PID:e245016;
                        PID:g1340088
      ##experimental_source strain H37Rv
GENETICS
   #gene        Rv1322
SUMMARY         #length 98  #molecular-weight 11334  #checksum 2740

  Query Match           72.1%;  Score 44;  DB 2;  Length 98;
  Best Local Similarity  66.7%;  Pred. No. 5.55e+00;
  Matches     6;  Conservative    0;  Mismatches   3;  Indels   0;  Gaps   0;

Db      24 EAGPDGHEY 32
           || | || |
Qy       1 EADPTGHSY 9


RESULT   14
ENTRY           B72600     #type complete
TITLE           hypothetical protein APE1266 - Aeropyrum pernix (strain K1)
ORGANISM        #formal_name Aeropyrum pernix
DATE            20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change
                  20-Aug-1999
ACCESSIONS      B72600
REFERENCE       A72450
   #authors     Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.;
                  Haikawa, Y.; Jin-no, K.; Takahashi, M.; Sekine, M.; Baba,
                  S.; Ankai, A.; Kosugi, H.; Hosoyama, A.; Fukui, S.; Nagai,
                  Y.; Nishijima, K.; Nakazawa, H.; Takamiya, M.; Masuda, S.;
                  Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;
                  Kushida, N.; Oguchi, A.; Aoki, K.; Kubota, K.; Nakamura,
                  Y.; Nomura, N.; Sako, Y.; Kikuchi, H.
   #journal     DNA Res. (1999) 6:83-101
   #title       Complete genome sequence of an aerobic hyper-thermophilic
                  Crenarchaeon, Aeropyrum pernix K1.
   #cross-references MUID:99310339
   #accession   B72600
      ##status        preliminary
      ##molecule_type DNA
      ##residues      1-128 ##label KAW
      ##cross-references DDBJ:AP000061; NID:g5104821; PIDN:BAA80256.1;
                        PID:d1044042; PID:g5104942
      ##experimental_source strain K1
GENETICS
   #gene        APE1266
SUMMARY         #length 128  #molecular-weight 13655  #checksum 5871

  Query Match           72.1%;  Score 44;  DB 2;  Length 128;
  Best Local Similarity  57.1%;  Pred. No. 5.55e+00;
  Matches     4;  Conservative    3;  Mismatches   0;  Indels   0;  Gaps   0;

Db      86 EPAGHGY 92
           :|:||:|
Qy       3 DPTGHSY 9


RESULT   15
ENTRY           T09933     #type complete
TITLE           nucleotide pyrophosphatase-like protein - Arabidopsis
                  thaliana
ALTERNATE_NAMES protein T16L4.210
ORGANISM        #formal_name Arabidopsis thaliana #common_name mouse-ear
                  cress
DATE            16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change
                  20-Sep-1999
ACCESSIONS      T09933
REFERENCE       Z16897
   #authors     Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Bancroft, I.;
                  Mewes, H.W.; Mayer, K.F.X.; Lemcke, K.; Schueller, C.
   #submission  submitted to the Protein Sequence Database, June 1999
   #accession   T09933
      ##status        preliminary
      ##molecule_type DNA
      ##residues      1-461 ##label BEV
      ##cross-references EMBL:AL079344; GSPDB:GN00062; ATSP:T16L4.210
      ##experimental_source cultivar Columbia; BAC clone T16L4
GENETICS
   #gene        ATSP:T16L4.210
   #map_position 4
SUMMARY         #length 461  #molecular-weight 51587  #checksum 161

  Query Match           72.1%;  Score 44;  DB 2;  Length 461;
  Best Local Similarity  55.6%;  Pred. No. 5.55e+00;
  Matches     5;  Conservative    3;  Mismatches   1;  Indels   0;  Gaps   0;

Db     219 EPDQSGHNY 227
           |:| :||:|
Qy       1 EADPTGHSY 9


Search completed: Tue Sep 12 13:19:15 2000
Job time : 9 secs.
```

****



****

Release 3.1A John F. Collins, Biocomputing Research Unit.

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Sep 12 13:18:14 2000; MasPar time 3.68 Seconds
75.785 Million cell updates/sec

Tabular output not generated.

Title: >US-08-819-669E-26
Description: (1-9) from US08819669E.pep
Perfect Score: 61
Sequence: 1 EADPTGHSY 9

Scoring table: PAM 150
Gap 15

Searched: 85661 seqs, 30989116 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot38
1:swissprot

Statistics: Mean 21.172; Variance 19.906; scale 1.064

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	61	100.0	309	1	MAG1_HUMAN	MELANOMA-ASSOCIATED AN	3.30e-05
2	54	88.5	234	1	MAG8_HUMAN	MELANOMA-ASSOCIATED AN	4.36e-03
3	54	88.5	315	1	MAG9_HUMAN	MELANOMA-ASSOCIATED AN	4.36e-03
4	52	85.2	369	1	MAGA_HUMAN	MELANOMA-ASSOCIATED AN	1.66e-02
5	51	83.6	319	1	MAGB_HUMAN	MELANOMA-ASSOCIATED AN	3.20e-02
6	48	78.7	268	1	RAD_RAT	GTP-BINDING PROTEIN RA	2.19e-01
7	48	78.7	269	1	RAD_HUMAN	GTP-BINDING PROTEIN RA	2.19e-01
8	48	78.7	873	1	PC1_HUMAN	PLASMA-CELL MEMBRANE G	2.19e-01
9	47	77.0	497	1	PEN3_ADE12	PENTON PROTEIN (VIRION	4.10e-01
10	46	75.4	314	1	MAG6_HUMAN	MELANOMA-ASSOCIATED AN	7.58e-01
11	46	75.4	346	1	MGB4_HUMAN	MELANOMA-ASSOCIATED AN	7.58e-01
12	45	73.8	503	1	VP57_BDV	57 KDA PROTEIN (P57).	1.39e+00
13	45	73.8	1033	1	T1R1_ECOLI	TYPE I RESTRICTION ENZ	1.39e+00
14	44	72.1	98	1	YD22_MYCTU	HYPOTHETICAL 11.3 KDA	2.53e+00
15	44	72.1	488	1	SUOX_HUMAN	SULFITE OXIDASE PRECUR	2.53e+00
16	44	72.1	488	1	SUOX_RAT	SULFITE OXIDASE PRECUR	2.53e+00
17	44	72.1	725	1	MYT1_HUMAN	MYELIN TRANSCRIPTION F	2.53e+00
18	43	70.5	290	1	YG1D_YEAST	HYPOTHETICAL 31.7 KDA	4.54e+00
19	43	70.5	314	1	MAG3_HUMAN	MELANOMA-ASSOCIATED AN	4.54e+00
20	43	70.5	347	1	MGB1_HUMAN	MELANOMA-ASSOCIATED AN	4.54e+00
21	43	70.5	506	1	Z157_HUMAN	ZINC FINGER PROTEIN 15	4.54e+00
22	43	70.5	669	1	COGU_HUMAN	MATRIX METALLOPROTEINA	4.54e+00
23	43	70.5	878	1	YB9X_YEAST	HYPOTHETICAL 98.1 KDA	4.54e+00
24	42	68.9	145	1	ANG3_MOUSE	ANGIOGENIN-3 PRECURSOR	8.07e+00
25	42	68.9	156	1	YLXS_BACSU	HYPOTHETICAL 17.6 KDA	8.07e+00
26	42	68.9	283	1	FOLD_BACSU	FOLD BIFUNCTIONAL PROT	8.07e+00
27	42	68.9	317	1	MAG4_HUMAN	MELANOMA-ASSOCIATED AN	8.07e+00
28	42	68.9	326	1	YKX2_CAEEL	HYPOTHETICAL 34.6 KDA	8.07e+00
29	42	68.9	347	1	PYRC_SALTY	DIHYDROOROTASE (EC 3.5	8.07e+00
30	42	68.9	392	1	CGL_CAEEL	PUTATIVE CYSTATHIONINE	8.07e+00
31	42	68.9	555	1	WETA_EMENI	REGULATORY PROTEIN WET	8.07e+00
32	42	68.9	747	1	GUND_CELFI	ENDOGLUCANASE D PRECUR	8.07e+00
33	42	68.9	3396	1	POLG_DEN1S	GENOME POLYPROTEIN [CO	8.07e+00
34	41	67.2	319	1	MGB2_HUMAN	MELANOMA-ASSOCIATED AN	1.42e+01
35	41	67.2	366	1	CD44_BOVIN	CD44 ANTIGEN PRECURSOR	1.42e+01
36	41	67.2	459	1	ZPR1_HUMAN	ZINC-FINGER PROTEIN ZP	1.42e+01
37	41	67.2	459	1	ZPR1_MOUSE	ZINC-FINGER PROTEIN ZP	1.42e+01
38	41	67.2	566	1	GUNG_CLOTM	ENDOGLUCANASE G PRECUR	1.42e+01
39	41	67.2	581	1	PBP2_NEIME	PENICILLIN-BINDING PRO	1.42e+01
40	41	67.2	581	1	PBP2_NEIGO	PENICILLIN-BINDING PRO	1.42e+01
41	41	67.2	1025	1	ITA8_HUMAN	INTEGRIN ALPHA-8.	1.42e+01
42	41	67.2	2411	1	DAB_DROME	DISABLED PROTEIN.	1.42e+01
43	41	67.2	3898	1	POLG_BVDVS	GENOME POLYPROTEIN.	1.42e+01
44	40	65.6	429	1	ACEA_RHOFA	ISOCITRATE LYASE (EC 4	2.47e+01
45	40	65.6	919	1	YKQ5_YEAST	HYPOTHETICAL 105.7 KDA	2.47e+01

ALIGNMENTS

RESULT 1
ID MAG1_HUMAN STANDARD; PRT; 309 AA.
AC P43355; O00346;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-FEB-2000 (Rel. 39, Last annotation update)
DE MELANOMA-ASSOCIATED ANTIGEN 1 (MAGE-1 ANTIGEN) (ANTIGEN MZ2-E).
GN MAGEA1 OR MAGE1 OR MAGE1A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 92086861.
RA van der Bruggen P., Traversari C., Chomez P., Lurquin C., de Plaen E.,
RA van den Eynde B., Knuth A., Boon T.;
RT "A gene encoding an antigen recognized by cytolytic T lymphocytes on
RT a human melanoma.";
RL Science 254:1643-1647(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=SKIN;
RX MEDLINE; 94311935.
RA Ding M., Beck R.J., Keller C.J., Fenton R.G.;
RT "Cloning and analysis of MAGE-1-related genes.";
RL Biochem. Biophys. Res. Commun. 202:549-555(1994).
RN [3]
RP SEQUENCE FROM N.A.
RA Gloeckner G., Rump A., Nordsiek G., Hinzmann B., Kioschis P.,
RA Heiss N., Poustka A., Bauer D., Drescher B., Knob A., Rosenthal A.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP MUTAGENESIS.
RC TISSUE=BLOOD;
RX MEDLINE; 94157413.
RA Gaugler B., van den Eynde B., van der Bruggen P., Romero P.,
RA Gaforio J.J., de Plaen E., Lethe B., Brasseur F., Boon T.;
RT "Human gene MAGE-3 codes for an antigen recognized on a melanoma by
RT autologous cytolytic T lymphocytes.";
RL J. Exp. Med. 179:921-930(1994).
RN [5]
RP SUBCELLULAR LOCATION.
RX MEDLINE; 95012905.
RA Schultz-Thater E., Juretic A., Dellabona P., Luscher U., Siegrist W.,
RA Harder F., Heberer M., Zuber M., Spagnoli G.C.;
RT "MAGE-1 gene product is a cytoplasmic protein.";
RL Int. J. Cancer 59:435-439(1994).

```
CC   -!- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL
CC       DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
CC       PROGRESSION. ANTIGEN RECOGNIZED ON A MELANOMA BY AUTOLOGOUS
CC       CYTOLYTIC T LYMPHOCYTES.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC       SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC       CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC       FOR TESTES. NEVER EXPRESSED IN KIDNEY TUMORS, LEUKEMIAS AND
CC       LYMPHOMAS.
CC   -!- POLYMORPHISM: THE VARIANT AT POSITION 32 LIKELY REPRESENTS A
CC       POLYMORPHISM OF THE MAG-1 GENE.
CC   -!- SIMILARITY: BELONGS TO THE MAGE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M77481; AAA03229.1; -.
DR   EMBL; U82672; AAB54061.1; -.
DR   MIM; 300016; -.
DR   PFAM; PF01454; MAGE; 1.
KW   Antigen; Multigene family; Polymorphism; Tumor antigen.
FT   VARIANT      32     32       T -> A.
FT                                /FTId=VAR_004283.
FT   DOMAIN       33     36       POLY-SER.
FT   MUTAGEN     163    163       D->A: ABOLISHES HLA-A1 BINDING.
FT   MUTAGEN     169    169       Y->A: ABOLISHES HLA-A1 BINDING.
FT   CONFLICT     72     72       R -> Q (IN REF. 3).
SQ   SEQUENCE   309 AA;  34342 MW;  544EEB1F9F4E9D33 CRC64;

  Query Match          100.0%;  Score 61;  DB 1;  Length 309;
  Best Local Similarity 100.0%;  Pred. No. 3.30e-05;
  Matches     9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db    161 EADPTGHSY 169
          |||||||||
Qy      1 EADPTGHSY 9


RESULT    2
ID   MAG8_HUMAN     STANDARD;      PRT;   234 AA.
AC   P43361;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   MELANOMA-ASSOCIATED ANTIGEN 8 (MAGE-8 ANTIGEN).
GN   MAGEA8 OR MAGE8.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95012457.
RA   de Plaen E., Arden K., Traversari C., Gaforio J.J., Szikora J.-P.,
RA   de Smet C., Brasseur F., van der Bruggen P., Lethe B., Lurquin C.,
RA   Brasseur R., Chomez P., de Backer O., Cavenee W., Boon T.;
RT   "Structure, chromosomal localization, and expression of 12 genes of
RT   the MAGE family.";
RL   Immunogenetics 40:360-369(1994).
CC   -!- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL
CC       DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
CC       PROGRESSION.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC       SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC       CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC       FOR TESTES AND PLACENTA.
CC   -!- SIMILARITY: BELONGS TO THE MAGE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U10693; AAA68876.1; -.
DR   PFAM; PF01454; MAGE; 1.
KW   Antigen; Multigene family; Tumor antigen.
FT   DOMAIN       40     43       POLY-SER.
SQ   SEQUENCE   234 AA;  25197 MW;  058A92EE6003A982 CRC64;

  Query Match          88.5%;  Score 54;  DB 1;  Length 234;
  Best Local Similarity 77.8%;  Pred. No. 4.36e-03;
  Matches     7;  Conservative    1;  Mismatches   1;  Indels   0;  Gaps   0;

Db    171 EVDPAGHSY 179
          | ||:||||
Qy      1 EADPTGHSY 9


RESULT    3
ID   MAG9_HUMAN     STANDARD;      PRT;   315 AA.
AC   P43362; Q92910;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   MELANOMA-ASSOCIATED ANTIGEN 9 (MAGE-9 ANTIGEN).
GN   MAGEA9 OR MAGE9.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95012457.
RA   de Plaen E., Arden K., Traversari C., Gaforio J.J., Szikora J.-P.,
RA   de Smet C., Brasseur F., van der Bruggen P., Lethe B., Lurquin C.,
RA   Brasseur R., Chomez P., de Backer O., Cavenee W., Boon T.;
RT   "Structure, chromosomal localization, and expression of 12 genes of
RT   the MAGE family.";
RL   Immunogenetics 40:360-369(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Timms K.M., Bondeson M.L., Ansari-Lari M.A., Lagerstedt K.,
RA   Nelson D.L., Pettersson U., Gibbs R.A.;
RL   Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL
CC       DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
CC       PROGRESSION.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC       SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC       CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC       FOR TESTES AND PLACENTA.
CC   -!- SIMILARITY: BELONGS TO THE MAGE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U10694; AAA68877.1; -.
DR   EMBL; U66083; AAB67888.1; -.
DR   PFAM; PF01454; MAGE; 1.
KW   Antigen; Multigene family; Tumor antigen.
FT   DOMAIN       34     37       POLY-GLU.
FT   DOMAIN       87     90       POLY-GLU.
SQ   SEQUENCE   315 AA;  35088 MW;  7FD2ED10D680D928 CRC64;
```

```
  Query Match           88.5%;  Score 54;  DB 1;  Length 315;
  Best Local Similarity  77.8%;  Pred. No. 4.36e-03;
  Matches    7;  Conservative    1;  Mismatches   1;  Indels   0;  Gaps   0;

Db    167 EVDPAGHSY 175
          | ||:||||
Qy      1 EADPTGHSY 9


RESULT    4
ID   MAGA_HUMAN     STANDARD;      PRT;   369 AA.
AC   P43363;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   MELANOMA-ASSOCIATED ANTIGEN 10 (MAGE-10 ANTIGEN).
GN   MAGEA10 OR MAGE10.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95012457.
RA   de Plaen E., Arden K., Traversari C., Gaforio J.J., Szikora J.-P.,
RA   de Smet C., Brasseur F., van der Bruggen P., Lethe B., Lurquin C.,
RA   Brasseur R., Chomez P., de Backer O., Cavenee W., Boon T.;
RT   "Structure, chromosomal localization, and expression of 12 genes of
RT   the MAGE family.";
RL   Immunogenetics 40:360-369(1994).
CC   -!- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL
CC       DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
CC       PROGRESSION.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC       SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC       CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC       FOR TESTES AND PLACENTA.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U10685; AAA68869.1; -.
DR   PFAM; PF01454; MAGE; 1.
KW   Antigen; Multigene family; Tumor antigen.
FT   DOMAIN       54     62       POLY-SER.
SQ   SEQUENCE   369 AA;  40766 MW;  16FA3301CAB716A6 CRC64;

  Query Match           85.2%;  Score 52;  DB 1;  Length 369;
  Best Local Similarity  77.8%;  Pred. No. 1.66e-02;
  Matches    7;  Conservative    1;  Mismatches   1;  Indels   0;  Gaps   0;

Db    193 EVDPTGHSF 201
          | ||||||:
Qy      1 EADPTGHSY 9


RESULT    5
ID   MAGB_HUMAN     STANDARD;      PRT;   319 AA.
AC   P43364;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   MELANOMA-ASSOCIATED ANTIGEN 11 (MAGE-11 ANTIGEN).
GN   MAGEA11 OR MAGE11.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95012457.
RA   de Plaen E., Arden K., Traversari C., Gaforio J.J., Szikora J.-P.,
RA   de Smet C., Brasseur F., van der Bruggen P., Lethe B., Lurquin C.,
RA   Brasseur R., Chomez P., de Backer O., Cavenee W., Boon T.;
RT   "Structure, chromosomal localization, and expression of 12 genes of
RT   the MAGE family.";
RL   Immunogenetics 40:360-369(1994).
CC   -!- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL
CC       DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
CC       PROGRESSION.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC       SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC       CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC       FOR TESTES AND PLACENTA.
CC   -!- SIMILARITY: BELONGS TO THE MAGE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U10686; AAA68870.1; -.
DR   PFAM; PF01454; MAGE; 1.
KW   Antigen; Multigene family; Tumor antigen.
SQ   SEQUENCE   319 AA;  35536 MW;  F51A0B4140277BE3 CRC64;

  Query Match           83.6%;  Score 51;  DB 1;  Length 319;
  Best Local Similarity  77.8%;  Pred. No. 3.20e-02;
  Matches    7;  Conservative    1;  Mismatches   1;  Indels   0;  Gaps   0;

Db    171 EVDPTSHSY 179
          | |||:|||
Qy      1 EADPTGHSY 9


RESULT    6
ID   RAD_RAT        STANDARD;      PRT;   268 AA.
AC   P55043;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   GTP-BINDING PROTEIN RAD (RAS ASSOCIATED WITH DIABETES) (RAD1).
GN   RRAD OR RAD.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LUNG;
RA   Rishi A.K., Gulamhussein A., Steele M.P.;
RL   Submitted (DEC-1994) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE RAD/GEM FAMILY OF GTP-BINDING
CC       PROTEINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U12187; AAA56719.1; -.
DR   HSSP; P10114; 2RAP.
DR   PFAM; PF00071; ras; 1.
KW   GTP-binding.
FT   NP_BIND      59     66       GTP (BY SIMILARITY).
FT   NP_BIND     108    112       GTP (BY SIMILARITY).
FT   NP_BIND     163    166       GTP (BY SIMILARITY).
SQ   SEQUENCE   268 AA;  29053 MW;  BF102CA24DD8090F CRC64;
```

```
  Query Match           78.7%;  Score 48;  DB 1;  Length 268;
  Best Local Similarity  55.6%;  Pred. No. 2.19e-01;
  Matches     5;  Conservative    4;  Mismatches   0;  Indels   0;  Gaps   0;

Db     80 EAEAAGHTY 88
          ||:::||:|
Qy      1 EADPTGHSY 9


RESULT    7
ID   RAD_HUMAN      STANDARD;      PRT;   269 AA.
AC   P55042;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   GTP-BINDING PROTEIN RAD (RAS ASSOCIATED WITH DIABETES) (RAD1).
GN   RRAD OR RAD.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=SKELETAL MUSCLE;
RX   MEDLINE; 94069319.
RA   Reynet C., Kahn C.R.;
RT   "Rad: a member of the Ras family overexpressed in muscle of type II
RT   diabetic humans.";
RL   Science 262:1441-1444(1993).
CC   -!- TISSUE SPECIFICITY: SKELETAL AND CARDIAC MUSCLE, LUNG, LESSER
CC       AMOUNTS IN PLACENTA AND KIDNEY. DETECTED IN ADIPOSE TISSUE.
CC       OVEREXPRESSED IN MUSCLE OF TYPE II DIABETIC HUMANS.
CC   -!- SIMILARITY: BELONGS TO THE RAD/GEM FAMILY OF GTP-BINDING
CC       PROTEINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L24564; AAA36540.1; -.
DR   HSSP; P10114; 2RAP.
DR   MIM; 179503; -.
DR   PFAM; PF00071; ras; 1.
KW   GTP-binding.
FT   NP_BIND      59     66       GTP (BY SIMILARITY).
FT   NP_BIND     108    112       GTP (BY SIMILARITY).
FT   NP_BIND     164    167       GTP (BY SIMILARITY).
SQ   SEQUENCE   269 AA;  29262 MW;  1802AEBE738A98BE CRC64;

  Query Match           78.7%;  Score 48;  DB 1;  Length 269;
  Best Local Similarity  55.6%;  Pred. No. 2.19e-01;
  Matches     5;  Conservative    4;  Mismatches   0;  Indels   0;  Gaps   0;

Db     80 EAEAAGHTY 88
          ||:::||:|
Qy      1 EADPTGHSY 9


RESULT    8
ID   PC1_HUMAN      STANDARD;      PRT;   873 AA.
AC   P22413;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   15-FEB-2000 (Rel. 39, Last annotation update)
DE   PLASMA-CELL MEMBRANE GLYCOPROTEIN PC-1 [INCLUDES: ALKALINE
DE   PHOSPHODIESTERASE I (EC 3.1.4.1); NUCLEOTIDE PYROPHOSPHATASE
DE   (EC 3.6.1.9) (NPPASE)].
GN   PDNP1 OR PC1 OR NPPS.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 91009202.
RA   Buckley M.F., Loveland K.A., McKinstry W.J., Garson O.M., Goding J.W.;
RT   "Plasma cell membrane glycoprotein PC-1. cDNA cloning of the human
RT   molecule, amino acid sequence, and chromosomal location.";
RL   J. Biol. Chem. 265:17506-17511(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 92246539.
RA   Funakoshi I., Kato H., Horie K., Yano T., Hori Y., Kobayashi H.,
RA   Inoue T., Suzuki H., Fukui S., Tsukahara M., Kajii T.,
RA   Yamashina I.;
RT   "Molecular cloning of cDNAs for human fibroblast nucleotide
RT   pyrophosphatase.";
RL   Arch. Biochem. Biophys. 295:180-187(1992).
CC   -!- FUNCTION: MAY HAVE A ROLE IN THE REGULATION OF N-GLYCOSYLATION.
CC   -!- CATALYTIC ACTIVITY: HYDROLYTICALLY REMOVES 5'-NUCLEOTIDES
CC       SUCCESSIVELY FROM THE 3'-HYDROXY TERMINI OF 3'-HYDROXY-TERMINATED
CC       OLIGO-NUCLEOTIDES.
CC   -!- CATALYTIC ACTIVITY: A DINUCLEOTIDE + H(2)O = 2 MONONUCLEOTIDE.
CC   -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
CC   -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN PLASMA CELLS AND ALSO IN A NUMBER
CC       OF NON-LYMPHOID TISSUES, INCLUDING THE DISTAL CONVOLUTED TUBULE
CC       OF THE KIDNEY, CHONDROCYTES, AND EPIDIDYMIS.
CC   -!- SIMILARITY: CONTAINS 2 SOMATOMEDIN-B TYPE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M57736; AAA63237.1; -.
DR   EMBL; D12485; BAA02054.1; -.
DR   EMBL; D12485; BAA02053.1; ALT_INIT.
DR   PIR; A39216; A39216.
DR   MIM; 173335; -.
DR   PFAM; PF01663; Phosphodiest; 1.
DR   PFAM; PF01033; Somatomedin_B; 2.
DR   PRINTS; PR00022; SOMATOMEDINB.
DR   PROSITE; PS00524; SOMATOMEDIN_B; 2.
KW   Glycoprotein; Transmembrane; Duplication; Signal-anchor; Hydrolase.
FT   DOMAIN        1     24       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     25     45       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT   DOMAIN       46    873       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN       52     92       SOMATOMEDIN-B LIKE.
FT   DOMAIN       93    136       SOMATOMEDIN-B LIKE.
FT   CARBOHYD    127    127       POTENTIAL.
FT   CARBOHYD    233    233       POTENTIAL.
FT   CARBOHYD    289    289       POTENTIAL.
FT   CARBOHYD    425    425       POTENTIAL.
FT   CARBOHYD    533    533       POTENTIAL.
FT   CARBOHYD    591    591       POTENTIAL.
FT   CARBOHYD    648    648       POTENTIAL.
FT   CARBOHYD    679    679       POTENTIAL.
FT   CARBOHYD    696    696       POTENTIAL.
SQ   SEQUENCE   873 AA;  99929 MW;  872808C20B048070 CRC64;

  Query Match           78.7%;  Score 48;  DB 1;  Length 873;
  Best Local Similarity  66.7%;  Pred. No. 2.19e-01;
  Matches     6;  Conservative    3;  Mismatches   0;  Indels   0;  Gaps   0;

Db    322 EPDSSGHSY 330
          |:|::||||
Qy      1 EADPTGHSY 9
```

```
RESULT    9
ID   PEN3_ADE12     STANDARD;      PRT;   497 AA.
AC   P36716;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   PENTON PROTEIN (VIRION COMPONENT III) (PENTON BASE PROTEIN).
GN   PIII.
OS   Human adenovirus type 12.
OC   Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 94076430.
RA   Sprengel J., Schmitz B., Heuss-Neitzel D., Zock C., Doerfler W.;
RT   "Nucleotide sequence of human adenovirus type 12 DNA: comparative
RT   functional analysis.";
RL   J. Virol. 68:379-389(1994).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X73487; CAA51887.1; -.
DR   PIR; S33938; S33938.
DR   PFAM; PF01686; Adeno_Penton_B; 1.
KW   Late protein.
SQ   SEQUENCE   497 AA;  56393 MW;  0524D989F5A9ED13 CRC64;

  Query Match           77.0%;  Score 47;  DB 1;  Length 497;
  Best Local Similarity  66.7%;  Pred. No. 4.10e-01;
  Matches     6;  Conservative    2;  Mismatches   1;  Indels   0;  Gaps   0;

Db    310 ETDPKGRSY 318
          |:|| |:||
Qy      1 EADPTGHSY 9


RESULT   10
ID   MAG6_HUMAN     STANDARD;      PRT;   314 AA.
AC   P43360;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   MELANOMA-ASSOCIATED ANTIGEN 6 (MAGE-6 ANTIGEN) (MAGE3B).
GN   MAGEA6 OR MAGE6.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95012457.
RA   de Plaen E., Arden K., Traversari C., Gaforio J.J., Szikora J.-P.,
RA   de Smet C., Brasseur F., van der Bruggen P., Lethe B., Lurquin C.,
RA   Brasseur R., Chomez P., de Backer O., Cavenee W., Boon T.;
RT   "Structure, chromosomal localization, and expression of 12 genes of
RT   the MAGE family.";
RL   Immunogenetics 40:360-369(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=SKIN;
RX   MEDLINE; 94311935.
RA   Ding M., Beck R.J., Keller C.J., Fenton R.G.;
RT   "Cloning and analysis of MAGE-1-related genes.";
RL   Biochem. Biophys. Res. Commun. 202:549-555(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95369706.
RA   Imai Y., Shichijo S., Yamada A., Katayama T., Yano H., Itoh K.;
RT   "Sequence analysis of the MAGE gene family encoding human tumor-
RT   rejection antigens.";
RL   Gene 160:287-290(1995).
CC   -!- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN TUMOR
CC       OR ASPECTS OF TUMOR PROGRESSION.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC       SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC       CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC       FOR TESTES.
CC   -!- SIMILARITY: BELONGS TO THE MAGE FAMILY. STRONG SIMILARITY TO
CC       MAGE-3.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U10691; AAA68875.1; -.
DR   EMBL; U10339; AAA19006.1; -.
DR   EMBL; D32076; BAA06842.1; -.
DR   MIM; 300176; -.
DR   PFAM; PF01454; MAGE; 1.
KW   Antigen; Multigene family; Tumor antigen.
FT   DOMAIN       40     43       POLY-SER.
SQ   SEQUENCE   314 AA;  34891 MW;  29B83C7FA6E50263 CRC64;

  Query Match           75.4%;  Score 46;  DB 1;  Length 314;
  Best Local Similarity  66.7%;  Pred. No. 7.58e-01;
  Matches     6;  Conservative    0;  Mismatches   3;  Indels   0;  Gaps   0;

Db    168 EVDPIGHVY 176
          | || || |
Qy      1 EADPTGHSY 9


RESULT   11
ID   MGB4_HUMAN     STANDARD;      PRT;   346 AA.
AC   O15481;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   MELANOMA-ASSOCIATED ANTIGEN B4 (MAGE-B4 ANTIGEN).
GN   MAGE-B4.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 98110575.
RA   Lurquin C., de Smet C., Brasseur F., Muscatelli F., Martelange V.,
RA   de Plaen E., Brasseur R., Monaco A.P., Boon T.;
RT   "Two members of the human MAGEB gene family located in Xp21.3 are
RT   expressed in tumors of various histological origins.";
RL   Genomics 46:397-408(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Muzny D., Arenson A.D., Adams C., Brundage E., Bunac C., Carvelli K.,
RA   Chacko J., Chen J., Di W., Ding Y., Dugan S., Durbin J., Forcum J.,
RA   Ganesh R., Garcia C., Goodman M., Gorrell J.H., Haywood M.,
RA   Hernandez J., Jackson L., Jin S., Kampal R., Karpathy S., Kovar C.,
RA   Leal B., Li Y., Lichtarge O., Liu W., Logan O., Lu J., Ly T.,
RA   Martinez C., Oswal G., Perez L., Rashid N.D., Rowland K., Savage L.,
RA   Scherer S.E., Shen H., Simon M., Stovall K., Timms K.M., Todd J.,
RA   Vo Q., Williamson A., Worley K.C., Yu W., Chinault C., Nelson D.,
RA   Gibbs R.A.;
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN TESTIS.
CC   -!- SIMILARITY: BELONGS TO THE MAGE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; U93163; AAC23619.1; -.
DR   EMBL; AC005185; AAD10637.1; -.
DR   MIM; 300153; -.
DR   PFAM; PF01454; MAGE; 1.
KW   Antigen; Multigene family.
SQ   SEQUENCE   346 AA;  38923 MW;  804F260BD50F036A CRC64;

  Query Match           75.4%;  Score 46;  DB 1;  Length 346;
  Best Local Similarity  66.7%;  Pred. No. 7.58e-01;
  Matches    6;  Conservative    1;  Mismatches   2;  Indels   0;  Gaps   0;

Db    168 EVNPTTHSY 176
          | :|| |||
Qy      1 EADPTGHSY 9


RESULT   12
ID   VP57_BDV       STANDARD;      PRT;   503 AA.
AC   P52638;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   57 KDA PROTEIN (P57).
OS   Borna disease virus (BDV).
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=V;
RX   MEDLINE; 94240137.
RA   Briese T., Schneemann A., Lewis A.J., Park Y.-S., Kim S.,
RA   Ludwig H., Lipkin W.I.;
RT   "Genomic organization of Borna disease virus.";
RL   Proc. Natl. Acad. Sci. U.S.A. 91:4362-4366(1994).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; U04608; AAA20227.1; -.
KW   Glycoprotein; Transmembrane.
FT   TRANSMEM      5     25       POTENTIAL.
FT   TRANSMEM    274    294       POTENTIAL.
FT   TRANSMEM    468    488       POTENTIAL.
FT   CARBOHYD     63     63       POTENTIAL.
FT   CARBOHYD    109    109       POTENTIAL.
FT   CARBOHYD    139    139       POTENTIAL.
FT   CARBOHYD    192    192       POTENTIAL.
FT   CARBOHYD    196    196       POTENTIAL.
FT   CARBOHYD    202    202       POTENTIAL.
FT   CARBOHYD    221    221       POTENTIAL.
FT   CARBOHYD    230    230       POTENTIAL.
FT   CARBOHYD    235    235       POTENTIAL.
FT   CARBOHYD    321    321       POTENTIAL.
FT   CARBOHYD    328    328       POTENTIAL.
FT   CARBOHYD    388    388       POTENTIAL.
FT   CARBOHYD    438    438       POTENTIAL.
SQ   SEQUENCE   503 AA;  56652 MW;  081B55347DF91A08 CRC64;

  Query Match           73.8%;  Score 45;  DB 1;  Length 503;
  Best Local Similarity  55.6%;  Pred. No. 1.39e+00;
  Matches    5;  Conservative    2;  Mismatches   2;  Indels   0;  Gaps   0;

Db    416 ETDPINHAY 424
          |:|| |:|
Qy      1 EADPTGHSY 9


RESULT   13
ID   T1R1_ECOLI     STANDARD;      PRT;  1033 AA.
AC   P10486;
DT   01-JUL-1989 (Rel. 11, Created)
DT   01-JUL-1989 (Rel. 11, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   TYPE I RESTRICTION ENZYME ECOR124II R PROTEIN (EC 3.1.21.3).
GN   HSDR OR HSR.
OS   Escherichia coli.
OG   Plasmid IncFIV R124/3.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 89178628.
RA   Price C., Lingner J., Bickle J., Firman T.A., Glover S.W.;
RT   "Basis for changes in DNA recognition by the EcoR124 and EcoR124/3
RT   type I DNA restriction and modification enzymes.";
RL   J. Mol. Biol. 205:115-125(1989).
CC   -!- FUNCTION: THE ECOR124/3 I ENZYME RECOGNIZES 5'GAA(N7)RTCG.
CC   -!- FUNCTION: SUBUNIT R IS REQUIRED FOR BOTH NUCLEASE AND ATPASE
CC       ACTIVITIES, BUT NOT FOR MODIFICATION.
CC   -!- SUBUNIT: THE TYPE I RESTRICTION & MODIFICATION SYSTEM IS COMPOSED
CC       OF THREE POLYPEPTIDES R,M AND S.
CC   -!- MISCELLANEOUS: TYPE I RESTRICTION AND MODIFICATION ENZYMES ARE
CC       COMPLEX, MULTIFUNCTIONAL SYSTEMS WHICH REQUIRE ATP, S-ADENOSYL
CC       METHIONINE AND MG(2+) AS CO-FACTORS AND, IN ADDITION TO THEIR
CC       ENDONUCLEOLYTIC AND METHYLASE ACTIVITIES, ARE POTENT DNA-DEPENDENT
CC       ATPASES.
CC   -!- SIMILARITY: WITH ATPASES.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X13145; CAA31543.1; -.
DR   PIR; S02168; S02168.
DR   REBASE; RB00989; EcoR124II.
KW   Plasmid; Restriction system; Hydrolase; DNA-binding; ATP-binding.
SQ   SEQUENCE   1033 AA;  119656 MW;  B55F3991356C1506 CRC64;

  Query Match           73.8%;  Score 45;  DB 1;  Length 1033;
  Best Local Similarity  75.0%;  Pred. No. 1.39e+00;
  Matches    6;  Conservative    1;  Mismatches   1;  Indels   0;  Gaps   0;

Db     23 AEPTGDSY 30
          |:||| ||
Qy      2 ADPTGHSY 9


RESULT   14
ID   YD22_MYCTU     STANDARD;      PRT;    98 AA.
AC   Q10635;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-FEB-2000 (Rel. 39, Last annotation update)
DE   HYPOTHETICAL 11.3 KDA PROTEIN RV1322.
GN   RV1322 OR MTCY130.07.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacter
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37RV;
```

```
RX   MEDLINE; 98295987.
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA   Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,
RA   Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA   Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA   Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA   Rutter S., Seeger K., Skelton S., Squares S., Sqares R., Sulston J.E.,
RA   Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the
RT   complete genome sequence.";
RL   Nature 393:537-544(1998).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z73902; CAA98086.1; -.
DR   TUBERCULIST; Rv1322; -.
KW   Hypothetical protein.
SQ   SEQUENCE   98 AA;  11334 MW;  72DF33A68405AE4B CRC64;

  Query Match           72.1%;  Score 44;  DB 1;  Length 98;
  Best Local Similarity  66.7%;  Pred. No. 2.53e+00;
  Matches     6;  Conservative    0;  Mismatches   3;  Indels   0;  Gaps   0;

Db     24 EAGPDGHEY 32
          || | || |
Qy      1 EADPTGHSY 9


RESULT   15
ID   SUOX_HUMAN     STANDARD;      PRT;   488 AA.
AC   P51687;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   SULFITE OXIDASE PRECURSOR (EC 1.8.3.1).
GN   SUOX.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE; 95322455.
RA   Garrett R.M., Bellissimo D.B., Rajagopalan K.V.;
RT   "Molecular cloning of human liver sulfite oxidase.";
RL   Biochim. Biophys. Acta 1262:147-149(1995).
RN   [2]
RP   VARIANTS GLN-160; ASP-208; TYR-370 AND ASP-473.
RX   MEDLINE; 98088796.
RA   Kisker C., Schindelin H., Pacheco A., Wehbi W.A., Garrett R.M.,
RA   Rajagopalan K.V., Enemark J.H., Rees D.C.;
RT   "Molecular basis of sulfite oxidase deficiency from the structure of
RT   sulfite oxidase.";
RL   Cell 91:973-983(1997).
RN   [3]
RP   VARIANT GLN-160.
RX   MEDLINE; 98263367.
RA   Garrett R.M., Johnson J.L., Graf T.N., Feigenbaum A.,
RA   Rajagopalan K.V.;
RT   "Human sulfite oxidase R160Q: identification of the mutation in a
RT   sulfite oxidase-deficient patient and expression and characterization
RT   of the mutant enzyme.";
RL   Proc. Natl. Acad. Sci. U.S.A. 95:6394-6398(1998).
CC   -!- CATALYTIC ACTIVITY: SULFITE + O(2) + H(2)O = SULFATE + H(2)O(2).
CC   -!- COFACTOR: MOLYBDENUM (MOLYBDOPTERIN) AND ONE PROTOHEME GROUP.
CC   -!- PATHWAY: TERMINAL REACTION IN THE OXIDATIVE DEGRADATION OF SULFUR-
CC       CONTAINING AMINO ACIDS. IT USES CYTOCHROME C AS AN ELECTRON
CC       ACCEPTOR.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INTER MEMBRANE SPACE.
CC   -!- DISEASE: DEFECTS IN SUOX ARE A CAUSE OF SULFITE OXIDASE
CC       DEFICIENCY; CHARACTERIZED BY NEUROLOGICAL ABNORMALITIES. OFTEN
CC       LEADS TO DEATH AT AN EARLY AGE.
CC   -!- SIMILARITY: WITH CYTOCHROME B5 AND NITRATE REDUCTASE.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L31573; AAA74886.1; -.
DR   HSSP; P07850; 1SOX.
DR   MIM; 272300; -.
DR   PFAM; PF00173; heme_1; 1.
DR   PFAM; PF00174; oxidored_molyb; 1.
DR   PRINTS; PR00407; EUMOPTERIN.
DR   PROSITE; PS00191; CYTOCHROME_B5_1; 1.
DR   PROSITE; PS00559; MOLYBDOPTERIN_EUK; 1.
KW   Oxidoreductase; Mitochondrion; Heme; Molybdenum; Transit peptide;
KW   Disease mutation.
FT   TRANSIT       1     22       MITOCHONDRION (BY SIMILARITY).
FT   CHAIN        23    488       SULFITE OXIDASE.
FT   DOMAIN       23    107       CYTOCHROME DOMAIN (BY SIMILARITY).
FT   DOMAIN      108    124       HINGE (BY SIMILARITY).
FT   DOMAIN      125    488       MOLYBDENUM-PTERIN DOMAIN (BY SIMILARITY).
FT   BINDING      61     61       HEME LIGAND (BY SIMILARITY).
FT   BINDING      86     86       HEME LIGAND (BY SIMILARITY).
FT   METAL       207    207       MOLYBDENUM-PTERIN (BY SIMILARITY).
FT   METAL       260    260       MOLYBDENUM-PTERIN (BY SIMILARITY).
FT   VARIANT     160    160       R -> Q (IN SUOX DEFICIENCY; 2% OF
FT                                ACTIVITY).
FT                                /FTId=VAR_002200.
FT   VARIANT     208    208       A -> D (IN SUOX DEFICIENCY).
FT                                /FTId=VAR_002201.
FT   VARIANT     370    370       S -> Y (IN SUOX DEFICIENCY).
FT                                /FTId=VAR_002202.
FT   VARIANT     473    473       G -> D (IN SUOX DEFICIENCY).
FT                                /FTId=VAR_002203.
SQ   SEQUENCE   488 AA;  53884 MW;  41EFA367FAB766DA CRC64;

  Query Match           72.1%;  Score 44;  DB 1;  Length 488;
  Best Local Similarity  55.6%;  Pred. No. 2.53e+00;
  Matches     5;  Conservative    3;  Mismatches   1;  Indels   0;  Gaps   0;

Db    265 DSDPTGTAY 273
          ::|||| :|
Qy      1 EADPTGHSY 9


Search completed: Tue Sep 12 13:18:21 2000
Job time : 7 secs.
```

****

MPSRCH (TM)

****

Release 3.1A John F. Collins, Biocomputing Research Unit.

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Sep 12 13:18:37 2000; MasPar time 8.87 Seconds
70.368 Million cell updates/sec

Tabular output not generated.

Title: >US-08-819-669E-26
Description: (1-9) from US08819669E.pep
Perfect Score: 61
Sequence: 1 EADPTGHSY 9

Scoring table: PAM 150
Gap 15

Searched: 225878 seqs, 69334122 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: sptrembl12
1:sp_archea 2:sp_bacteria 3:sp_fungi 4:sp_human
5:sp_invertebrate 6:sp_mammal 7:sp_mhc 8:sp_organelle
9:sp_phage 10:sp_plant 11:sp_rodent 12:sp_unclassified
13:sp_vertebrate 14:sp_virus

Statistics: Mean 20.592; Variance 19.613; scale 1.050

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	49	80.3	347	4	O00601	DAM10=DSS-AHC CRITICAL	2.04e-01
2	49	80.3	347	4	O75862	MAGE-B1.	2.04e-01
3	48	78.7	129	1	Q9YDL2	129AA LONG HYPOTHETICA	3.86e-01
4	48	78.7	308	11	O88667	RAS-LIKE GTP-BINDING P	3.86e-01
5	48	78.7	308	4	Q92788	RAD GTPASE.	3.86e-01
6	47	77.0	3942	11	O88737	BASSOON.	7.24e-01
7	46	75.4	1187	2	Q59278	ENDOXYLANASE (EC 3.2.1	1.35e+00
8	45	73.8	131	2	O85701	HYPOTHETICAL 14.2 KD P	2.48e+00
9	45	73.8	330	11	Q60763	MELANOMA ANTIGEN, RELA	2.48e+00
10	45	73.8	330	11	Q60761	MELANOMA ANTIGEN RELAT	2.48e+00
11	45	73.8	353	14	Q88626	ORFIV.	2.48e+00
12	45	73.8	503	14	Q9WNA0	GLYCOPROTEIN GP94.	2.48e+00
13	45	73.8	503	14	O10403	P57 (FRAGMENT).	2.48e+00
14	45	73.8	503	14	O10399	P57 (FRAGMENT).	2.48e+00
15	45	73.8	503	14	O10394	P57 (FRAGMENT).	2.48e+00
16	45	73.8	503	14	Q65459	GLYCOPROTEIN.	2.48e+00
17	45	73.8	503	14	O10400	P57 (FRAGMENT).	2.48e+00
18	45	73.8	503	14	Q89857	GLYCOPROTEIN.	2.48e+00
19	45	73.8	503	14	O10397	P57 (FRAGMENT).	2.48e+00
20	45	73.8	1032	11	Q61989	INTEGRIN ALPHA-4 SUBUN	2.48e+00
21	44	72.1	128	1	Q9YCJ1	128AA LONG HYPOTHETICA	4.54e+00
22	44	72.1	320	11	O89006	MAGEA1 PROTEIN.	4.54e+00
23	44	72.1	325	11	O89010	MAGEA6 PROTEIN.	4.54e+00
24	44	72.1	1078	11	O08995	MYELIN TRANSCRIPTION F	4.54e+00
25	44	72.1	1121	4	O94922	KIAA0835 PROTEIN.	4.54e+00
26	44	72.1	2457	14	O41965	TEGUMENT PROTEIN.	4.54e+00
27	44	72.1	3938	11	O88778	BASSOON.	4.54e+00
28	43	70.5	126	6	Q9XSI8	MT2-MMP PROTEIN (FRAGM	8.20e+00
29	43	70.5	156	5	Q23961	ALPHA-AMYLASE 2 (EC 3.	8.20e+00
30	43	70.5	197	2	O53701	HYPOTHETICAL 21.4 KD P	8.20e+00
31	43	70.5	272	11	O35830	LECITHIN-CHOLESTEROL A	8.20e+00
32	43	70.5	320	11	O89009	MAGEA5 PROTEIN.	8.20e+00
33	43	70.5	341	5	Q25028	CYSTEINE PROTEINASE.	8.20e+00
34	43	70.5	453	4	Q9Y6X5	KIAA0879 PROTEIN.	8.20e+00
35	43	70.5	564	4	Q14111	MATRIX METALLOPROTEINA	8.20e+00
36	43	70.5	657	11	O54732	MATRIX METALLOPROTEINA	8.20e+00
37	43	70.5	668	5	Q9XXU5	B0019.1 PROTEIN.	8.20e+00
38	43	70.5	745	2	O86648	HYPOTHETICAL 81.0 KD P	8.20e+00
39	43	70.5	875	9	O64044	YOMG.	8.20e+00
40	43	70.5	875	2	O31978	YOMG PROTEIN.	8.20e+00
41	43	70.5	1086	2	O69230	ENDO-1,4-BETA-XYLANASE	8.20e+00
42	43	70.5	3851	4	O43161	BASSOON PROTEIN (KIAA0	8.20e+00
43	42	68.9	503	2	O33360	HYPOTHETICAL 54.2 KD P	1.47e+01
44	42	68.9	503	2	O33266	HYPOTHETICAL 54.1 KD P	1.47e+01
45	42	68.9	568	10	O80487	T12M4.10 PROTEIN.	1.47e+01

ALIGNMENTS

RESULT 1
ID O00601 PRELIMINARY; PRT; 347 AA.
AC O00601;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE DAM10=DSS-AHC CRITICAL INTERVAL MAGE SUPERFAMILY PROTEIN.
GN DAM10.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RX MEDLINE; 96081328.
RA DABOVIC B., ZANARIA E., BARDONI B., LISA A., BORDIGNON C., RUSSO V.,
RA MATESSI C., TRAVERSARI C., CAMERINO G.;
RT "A family of rapidly evolving genes from the sex reversal critical
RT region in Xp21.";
RL Mamm. Genome 6:571-580(1995).
DR EMBL; S80936; AAC97145.1; -.
DR PFAM; PF01454; MAGE; 1.
SQ SEQUENCE 347 AA; 39049 MW; AB96D5BB CRC32;

Query Match 80.3%; Score 49; DB 4; Length 347;
Best Local Similarity 55.6%; Pred. No. 2.04e-01;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```
Db    167 EDNPSGHTY 175
          | :|:||:|
Qy      1 EADPTGHSY 9
```

RESULT 2
ID O75862 PRELIMINARY; PRT; 347 AA.
AC O75862;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE MAGE-B1.
GN MAGE-B1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.

```
RN   [1]
RP   SEQUENCE FROM N.A.
RA   MUZNY D., ARENSON A.D., ADAMS C., BRUNDAGE E., BUNAC C., CARVELLI K.,
RA   CHACKO J., CHEN J., DI W., DING Y., DUGAN S., DURBIN J., FORCUM J.,
RA   GANESH R., GARCIA C., GOODMAN M., GORRELL J.H., HAYWOOD M.,
RA   HERNANDEZ J., JACKSON L., JIN S., KAMPAL R., KARPATHY S., KOVAR C.,
RA   LEAL B., LI Y., LICHTARGE O., LIU W., LOGAN O., LU J., LY T.,
RA   MARTINEZ C., OSWAL G., PEREZ L., RASHID N.D., ROWLAND K., SAVAGE L.,
RA   SCHERER S.E., SHEN H., SIMON M., STOVALL K., TIMMS K.M., TODD J.,
RA   VO Q., WILLIAMSON A., WORLEY K.C., YU W., CHINAULT C., NELSON D.,
RA   GIBBS R.A.;
RT   "Direct Submission.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC005185; AAD10634.1; -.
DR   PFAM; PF01454; MAGE; 1.
SQ   SEQUENCE   347 AA;  39037 MW;  48A34904 CRC32;

  Query Match           80.3%;  Score 49;  DB 4;  Length 347;
  Best Local Similarity  55.6%;  Pred. No. 2.04e-01;
  Matches    5;  Conservative    3;  Mismatches   1;  Indels   0;  Gaps   0;

Db    167 EDNPSGHTY 175
          | :|:||:|
Qy      1 EADPTGHSY 9


RESULT    3
ID   Q9YDL2      PRELIMINARY;      PRT;   129 AA.
AC   Q9YDL2;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   129AA LONG HYPOTHETICAL PROTEIN.
GN   APE0901.
OS   Aeropyrum pernix.
OC   Archaea; Crenarchaeota; Aeropyrum.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K1;
RX   MEDLINE; 99310339.
RA   KAWARABAYASI Y., HINO Y., HORIKAWA H., YAMAZAKI S., HAIKAWA Y.,
RA   JIN-NO K., TAKAHASHI M., SEKINE M., BABA S., ANKAI A., KOSUGI H.,
RA   HOSOYAMA A., FUKUI S., NAGAI Y., NISHIJIMA K., NAKAZAWA H.,
RA   TAKAMIYA M., MASUDA S., FUNAHASHI T., TANAKA T., KUDOH Y.,
RA   YAMAZAKI J., KUSHIDA N., OGUCHI A., AOKI K., KUBOTA K., NAKAMURA Y.,
RA   NOMURA N., SAKO Y., KIKUCHI H.;
RT   "Complete genome sequence of an aerobic hyper-thermophilic
RT   crenarchaeon, Aeropyrum pernix K1.";
RL   DNA Res. 6:83-101(1999).
DR   EMBL; AP000060; BAA79885.1; -.
SQ   SEQUENCE   129 AA;  14303 MW;  A2EB2774 CRC32;

  Query Match           78.7%;  Score 48;  DB 1;  Length 129;
  Best Local Similarity  85.7%;  Pred. No. 3.86e-01;
  Matches    6;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Db    115 DPAGHSY 121
          ||:||||
Qy      3 DPTGHSY 9


RESULT    4
ID   O88667      PRELIMINARY;      PRT;   308 AA.
AC   O88667;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   RAS-LIKE GTP-BINDING PROTEIN RAD.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   FINLIN B.S., ANDRES D.A.;
RT   "Cloning of the mouse Rad gene.";
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF084466; AAC33133.1; -.
DR   HSSP; P10114; 2RAP.
DR   PFAM; PF00071; ras; 1.
SQ   SEQUENCE   308 AA;  33279 MW;  10AE9F3C CRC32;

  Query Match           78.7%;  Score 48;  DB 11;  Length 308;
  Best Local Similarity  55.6%;  Pred. No. 3.86e-01;
  Matches    5;  Conservative    4;  Mismatches   0;  Indels   0;  Gaps   0;

Db    119 EAEAAGHTY 127
          ||:::||:|
Qy      1 EADPTGHSY 9


RESULT    5
ID   Q92788      PRELIMINARY;      PRT;   308 AA.
AC   Q92788;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   RAD GTPASE.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 94069319.
RA   REYNET C., KAHN C.R.;
RT   "Rad: a member of the Ras family overexpressed in muscle of type II
RT   diabetic humans.";
RL   Science 262:1441-1444(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 96375161.
RA   CALDWELL J.S., MOYERS J.S., DORIA A., REYNET C., KAHN R.C.;
RT   "Molecular cloning of the human rad gene: gene structure and complete
RT   nucleotide sequence.";
RL   Biochim. Biophys. Acta 1316:145-148(1996).
DR   EMBL; U46165; AAB17064.1; -.
DR   HSSP; P10114; 2RAP.
DR   PFAM; PF00071; ras; 1.
SQ   SEQUENCE   308 AA;  33220 MW;  4B7A3673 CRC32;

  Query Match           78.7%;  Score 48;  DB 4;  Length 308;
  Best Local Similarity  55.6%;  Pred. No. 3.86e-01;
  Matches    5;  Conservative    4;  Mismatches   0;  Indels   0;  Gaps   0;

Db    119 EAEAAGHTY 127
          ||:::||:|
Qy      1 EADPTGHSY 9


RESULT    6
ID   O88737      PRELIMINARY;      PRT;  3942 AA.
AC   O88737;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   BASSOON.
GN   BASSOON.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129 SVJ;
RX   MEDLINE; 98345363.
RA   DIECK S., SANMARTI-VILA L., LANGNAESE K., RICHTER K., KINDLER S.,
RA   SOYKE A., WEX H., SMALLA K.H., KAMPF U., FRANZER J.T., STUMM M.,
RA   GARNER C.C., GUNDELFINGER E.D.;
```

```
RT   "Bassoon, a novel zinc-finger CAG/glutamine-repeat protein selectively
RT   localized at the active zone of presynaptic nerve terminals.";
RL   J. Cell Biol. 142:499-509(1998).
DR   EMBL; Y17034; CAA76598.1; -.
DR   EMBL; Y17035; CAA76598.1; JOINED.
DR   EMBL; Y17036; CAA76598.1; JOINED.
DR   EMBL; Y17037; CAA76598.1; JOINED.
DR   EMBL; Y17038; CAA76598.1; JOINED.
SQ   SEQUENCE   3942 AA;  418739 MW;  9D6C5BC6 CRC32;

  Query Match           77.0%;  Score 47;  DB 11;  Length 3942;
  Best Local Similarity  55.6%;  Pred. No. 7.24e-01;
  Matches    5;  Conservative    4;  Mismatches   0;  Indels   0;  Gaps   0;

Db   1589 DSQPTSHSY 1597
          :::||:|||
Qy      1 EADPTGHSY 9


RESULT    7
ID   Q59278      PRELIMINARY;      PRT;  1187 AA.
AC   Q59278;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   ENDOXYLANASE (EC 3.2.1.8) (ENDO-1,4-BETA-XYLANASE)
DE   (1,4-BETA-D-XYLAN XYLANOHYDROLASE).
GN   XYNC.
OS   Cellulomonas fimi.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Micrococcineae; Cellulomonadaceae; Cellulomonas.
RN   [1]
RP   SEQUENCE OF 1-352 FROM N.A.
RX   MEDLINE; 96245431.
RA   CLARKE J.H., DAVIDSON K., GILBERT H.J., FONTES C.M., HAZLEWOOD G.P.;
RT   "A modular xylanase from mesophilic Cellulomonas fimi contains the
RT   same cellulose-binding and thermostabilizing domains as xylanases from
RT   thermophilic bacteria.";
RL   FEMS Microbiol. Lett. 139:27-35(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   CLARKE J.H.;
RL   Submitted (AUG-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-XYLOSIDIC
CC       LINKAGES IN XYLANS.
DR   EMBL; Z50866; CAA90745.1; -.
DR   HSSP; P14768; 1CLX.
DR   PFAM; PF00331; Glyco_hydro_10; 1.
DR   PFAM; PF01522; Polysac_deacet; 1.
DR   PRINTS; PR00134; GLHYDRLASE10.
KW   Xylan degradation; Hydrolase; Glycosidase.
SQ   SEQUENCE   1187 AA;  125378 MW;  92B3994A CRC32;

  Query Match           75.4%;  Score 46;  DB 2;  Length 1187;
  Best Local Similarity  75.0%;  Pred. No. 1.35e+00;
  Matches    6;  Conservative    2;  Mismatches   0;  Indels   0;  Gaps   0;

Db   1046 TDPTGRSY 1053
          :||||:||
Qy      2 ADPTGHSY 9


RESULT    8
ID   O85701      PRELIMINARY;      PRT;  131 AA.
AC   O85701;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   HYPOTHETICAL 14.2 KD PROTEIN.
OS   Streptomyces lividans.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1326, AJ100;
RA   ALTENBUCHNER J.;
RT   "Amplifiable element AUD4 from Streptomyces lividans 66.";
RL   Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF072709; AAC25770.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   131 AA;  14187 MW;  8321BCE1 CRC32;

  Query Match           73.8%;  Score 45;  DB 2;  Length 131;
  Best Local Similarity  62.5%;  Pred. No. 2.48e+00;
  Matches    5;  Conservative    3;  Mismatches   0;  Indels   0;  Gaps   0;

Db   110 SDPAGHSF 117
          :||:|||:
Qy      2 ADPTGHSY 9


RESULT    9
ID   Q60763      PRELIMINARY;      PRT;  330 AA.
AC   Q60763;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   MELANOMA ANTIGEN, RELATED SEQUENCE 2 (SMAGE-3 PROTEIN).
GN   MAGE-RS3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DBA/2; TISSUE=KIDNEY;
RX   MEDLINE; 96070435.
RA   DE BACKER O., VERHEYDEN A.M., MARTIN B., GODELAINE D., DE PLAEN E.,
RA   BRASSEUR R., AVNER P., BOON T.;
RT   "Structure, chromosomal location, and expression pattern of three
RT   mouse genes homologous to the human MAGE genes.";
RL   Genomics 28:74-83(1995).
DR   EMBL; U19033; AAA86098.1; -.
DR   PFAM; PF01454; MAGE; 1.
SQ   SEQUENCE   330 AA;  35985 MW;  83AD4246 CRC32;

  Query Match           73.8%;  Score 45;  DB 11;  Length 330;
  Best Local Similarity  66.7%;  Pred. No. 2.48e+00;
  Matches    6;  Conservative    1;  Mismatches   2;  Indels   0;  Gaps   0;

Db   161 EIDPSTHSY 169
          | ||: |||
Qy      1 EADPTGHSY 9


RESULT   10
ID   Q60761      PRELIMINARY;      PRT;  330 AA.
AC   Q60761; Q60762;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   MELANOMA ANTIGEN RELATED SEQUENCES 1 AND 2
DE   (SMAGE-1 PROTEIN / SMAGE-2 PROTEIN).
GN   MAGEB2 OR (MAGE-RS1 OR SMAGE1) AND (MAGE-RS2 OR SMAGE2).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC   Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DBA/2; TISSUE=KIDNEY;
RX   MEDLINE; 96070435.
RA   DE BACKER O., VERHEYDEN A.-M., MARTIN B., GODELAINE D., DE PLAEN E.,
RA   BRASSEUR R., AVNER P., BOON T.;
RT   "Structure, chromosomal location, and expression pattern of three
RT   mouse genes homologous to the human MAGE genes.";
RL   Genomics 28:74-83(1995).
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN TUMOURS OF VARIOUS HISTOLOGICAL
```

```
CC       TYPES BUT NOT IN NORMAL TISSUES EXCEPT TESTIS.
CC   -!- SIMILARITY: BELONGS TO THE MAGE FAMILY.
DR   EMBL; U19031; AAA86096.1; ALT_INIT.
DR   EMBL; U19032; AAA86097.1; -.
DR   MGD; MGI:105117; Mageb2.
DR   PFAM; PF01454; MAGE; 1.
KW   Antigen; Tumor antigen.
SQ   SEQUENCE   330 AA;  35936 MW;  36D760C5 CRC32;

  Query Match           73.8%;  Score 45;  DB 11;  Length 330;
  Best Local Similarity  66.7%;  Pred. No. 2.48e+00;
  Matches    6;  Conservative    1;  Mismatches   2;  Indels   0;  Gaps   0;

Db    161 EIDPSTHSY 169
          | ||: |||
Qy      1 EADPTGHSY 9


RESULT   11
ID   Q88626      PRELIMINARY;      PRT;   353 AA.
AC   Q88626;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   ORFIV.
OS   Borna disease virus (BDV).
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 94149825.
RA   CUBITT B., OLDSTONE C., LA TORRE J.;
RT   "Sequence and genome organization of Borna disease virus.";
RL   J. Virol. 68:1382-1396(1994).
DR   EMBL; L27077; AAA20666.1; -.
SQ   SEQUENCE   353 AA;  39959 MW;  555715F0 CRC32;

  Query Match           73.8%;  Score 45;  DB 14;  Length 353;
  Best Local Similarity  55.6%;  Pred. No. 2.48e+00;
  Matches    5;  Conservative    2;  Mismatches   2;  Indels   0;  Gaps   0;

Db    266 ETDPINHAY 274
          |:|| |:|
Qy      1 EADPTGHSY 9


RESULT   12
ID   Q9WNA0      PRELIMINARY;      PRT;   503 AA.
AC   Q9WNA0;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   GLYCOPROTEIN GP94.
OS   Borna disease virus (BDV).
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=RW98;
RX   MEDLINE; 99329142.
RA   PLANZ O., RENTZSCH C., BATRA A., BATRA A., WINKLER T., BUETTNER M.,
RA   RZIHA H.-J., STITZ L.;
RT   "Pathogenesis of borna disease virus: granulocyte fractions of
RT   psychiatric patients harbor infectious virus in the absence of
RT   antiviral antibodies.";
RL   J. Virol. 73:6251-6256(1999).
DR   EMBL; AF158633; AAD45291.1; -.
SQ   SEQUENCE   503 AA;  56588 MW;  EC993A56 CRC32;

  Query Match           73.8%;  Score 45;  DB 14;  Length 503;
  Best Local Similarity  55.6%;  Pred. No. 2.48e+00;
  Matches    5;  Conservative    2;  Mismatches   2;  Indels   0;  Gaps   0;

Db    416 ETDPINHAY 424
          |:|| |:|
Qy      1 EADPTGHSY 9


RESULT   13
ID   O10403      PRELIMINARY;      PRT;   503 AA.
AC   O10403;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-AUG-1998 (TrEMBLrel. 07, Last annotation update)
DE   P57 (FRAGMENT).
OS   Borna disease virus (BDV).
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   ZIMMERMANN W., KOKORSCH J., DUERRWALD R., LUDWIG H.;
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U94882; AAB53731.1; -.
FT   NON_TER     503    503
SQ   SEQUENCE   503 AA;  56510 MW;  E59DC9A6 CRC32;

  Query Match           73.8%;  Score 45;  DB 14;  Length 503;
  Best Local Similarity  55.6%;  Pred. No. 2.48e+00;
  Matches    5;  Conservative    2;  Mismatches   2;  Indels   0;  Gaps   0;

Db    416 ETDPINHAY 424
          |:|| |:|
Qy      1 EADPTGHSY 9


RESULT   14
ID   O10399      PRELIMINARY;      PRT;   503 AA.
AC   O10399;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-AUG-1998 (TrEMBLrel. 07, Last annotation update)
DE   P57 (FRAGMENT).
OS   Borna disease virus (BDV).
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   ZIMMERMANN W., KOKORSCH J., DUERRWALD R., LUDWIG H.;
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U94874; AAB53723.1; -.
FT   NON_TER     503    503
SQ   SEQUENCE   503 AA;  56564 MW;  28E55C1E CRC32;

  Query Match           73.8%;  Score 45;  DB 14;  Length 503;
  Best Local Similarity  55.6%;  Pred. No. 2.48e+00;
  Matches    5;  Conservative    2;  Mismatches   2;  Indels   0;  Gaps   0;

Db    416 ETDPINHAY 424
          |:|| |:|
Qy      1 EADPTGHSY 9


RESULT   15
ID   O10394      PRELIMINARY;      PRT;   503 AA.
AC   O10394;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-AUG-1998 (TrEMBLrel. 07, Last annotation update)
DE   P57 (FRAGMENT).
OS   Borna disease virus (BDV).
OC   Viruses; ssRNA negative-strand viruses; Mononegavirales.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=436;
RA   ZIMMERMANN W., KOKORSCH J., LUNDGREN A.L., LUDWIG H.;
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U94866; AAB53715.1; -.
FT   NON_TER     503    503
SQ   SEQUENCE   503 AA;  56578 MW;  B543AFC4 CRC32;
```

Query Match 73.8%; Score 45; DB 14; Length 503;
Best Local Similarity 55.6%; Pred. No. 2.48e+00;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```
Db    416 ETDPINHAY 424
          |:||  |:|
Qy      1 EADPTGHSY 9
```

Search completed: Tue Sep 12 13:18:50 2000
Job time : 13 secs.

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Sep 13 07:14:17 2000; MasPar time 3.65 Seconds
58.333 Million cell updates/sec

Tabular output not generated.

Title: >US-08-819-669E-26
Description: (1-9) from US08819669E.pep
Perfect Score: 61
Sequence: 1 EADPTGHSY 9

Scoring table: PAM 150
Gap 15

Searched: 188963 seqs, 23686106 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq36
1:geneseqp

Statistics: Mean 15.425; Variance 35.537; scale 0.434

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	61	100.0	9	1	W98945	HLA-A1 binding peptide	2.32e-01
2	61	100.0	9	1	Y01727	Exemplary antigenic pe	2.32e-01
3	61	100.0	9	1	Y10633	Peptide antigen SEQ ID	2.32e-01
4	61	100.0	9	1	Y00685	Tumour antigen booster	2.32e-01
5	61	100.0	9	1	Y10424	HLA Class I motif pept	2.32e-01
6	61	100.0	9	1	Y10623	Peptide antigen SEQ ID	2.32e-01
7	61	100.0	9	1	W54622	Peptide from Mage-1 16	2.32e-01
8	61	100.0	9	1	R50281	MAGE-1 nonapeptide.	2.32e-01
9	61	100.0	9	1	R83932	MHC class I restricted	2.32e-01
10	61	100.0	9	1	R29769	Antigen E peptide.	2.32e-01
11	61	100.0	9	1	W68371	Human MAGE-1 peptide b	2.32e-01
12	61	100.0	9	1	W77125	gp75/TRP-1 synthetic p	2.32e-01
13	61	100.0	9	1	W75734	Peptidase-resistant pe	2.32e-01
14	61	100.0	9	1	W00897	Human melanoma MAGE1 t	2.32e-01
15	61	100.0	9	1	W75736	Peptidase-resistant pe	2.32e-01
16	61	100.0	9	1	R82988	P815 antigenic peptide	2.32e-01
17	61	100.0	9	1	R90692	Human leukocyte antige	2.32e-01
18	61	100.0	9	1	W56729	MAGE-1 antigenic parti	2.32e-01
19	61	100.0	9	1	R75954	Melanoma antigen (MAGE	2.32e-01
20	61	100.0	9	1	R99343	MAGE-1 nonapeptide.	2.32e-01
21	61	100.0	9	1	W78838	MAGE-1 protein fragmen	2.32e-01
22	61	100.0	9	1	R65112	MAGE 1 immunogenic pep	2.32e-01
23	61	100.0	9	1	R63675	Synthetic peptide deri	2.32e-01
24	61	100.0	9	1	R78824	MAGE-1 cytotoxic T lym	2.32e-01
25	61	100.0	9	1	R65135	MAGE 1 immunogenic pep	2.32e-01
26	61	100.0	9	1	R49224	HLA-A1 MAGE 1 antigen	2.32e-01
27	61	100.0	9	1	R47330	HLA-A1 MAGE 1 antigen	2.32e-01
28	61	100.0	10	1	W23038	MAGE-1/HLA-B44 tumour	2.32e-01
29	61	100.0	12	1	R80620	Immunogenic peptide of	2.32e-01
30	61	100.0	309	1	R70909	Human melanoma antigen	2.32e-01
31	61	100.0	309	1	W81548	Tumour rejection antig	2.32e-01
32	59	96.7	9	1	R99342	HLA binding nonapeptid	4.62e-01
33	58	95.1	9	1	R99339	HLA binding nonapeptid	6.49e-01
34	57	93.4	9	1	W75733	Peptidase-resistant pe	9.12e-01
35	57	93.4	9	1	W75735	Peptidase-resistant pe	9.12e-01
36	55	90.2	9	1	R99337	HLA binding nonapeptid	1.79e+00
37	55	90.2	9	1	R99340	HLA binding nonapeptid	1.79e+00
38	54	88.5	9	1	R99338	HLA binding nonapeptid	2.50e+00
39	52	85.2	9	1	Y10628	Peptide antigen SEQ ID	4.86e+00
40	49	80.3	9	1	R99341	HLA binding nonapeptid	1.30e+01
41	48	78.7	308	1	R45431	Diabetogene rad: A typ	1.79e+01
42	48	78.7	308	1	W13869	Rad protein.	1.79e+01
43	48	78.7	925	1	R79148	Human insulin receptor	1.79e+01
44	47	77.0	9	1	Y10629	Peptide antigen SEQ ID	2.47e+01
45	47	77.0	272	1	W44186	Maleate cis-trans isom	2.47e+01

ALIGNMENTS

RESULT 1
ID W98945 standard; peptide; 9 AA.
AC W98945;
DT 06-MAY-1999 (first entry)
DE HLA-A1 binding peptide derived from MAGE-1.
KW Human leukocyte antigen; HLA; HLA-A2 binding peptide; T cell;
KW cytolytic T cell; CTL.
OS Synthetic.
OS Homo sapiens.
PN WO9858951-A1.
PD 30-DEC-1998.
PF 18-JUN-1998; U12879.
PR 16-APR-1998; US-061388.
PR 23-JUN-1997; US-880963.
PA (LUDW-) LUDWIG INST CANCER RES.
PI Cerottini J, Romero P, Valmori D;
DR WPI; 99-105609/09.
PT New decamer peptides which bind to HLA molecules - useful to
PT identify HLA-A2 positive cells and provoke T cells
PS Example 7; Page 18; 45pp; English.
CC The present invention describes peptides which bind to an HLA-A2
CC molecule and have Val at the carboxy terminus, and either: (a) Ala, Tyr
CC or Phe at the amino terminus, and Ala at position 2 (P1); or (b) Glu at
CC the amino terminus, and Ala, Leu, or Met at positions 2 and 3, with the
CC proviso that Ala is not at both positions (P2). The peptides of the
CC present invention are used to identify HLA-A2 positive cells, provoke
CC T cells, and determine the presence of particular T cells including
CC cytolytic T cells (CTLs). They provide a better target than the prior
CC art CTL-stimulating peptide. The present sequence represents a peptide
CC used in an example from the present invention.
SQ Sequence 9 AA;

Query Match 100.0%; Score 61; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.32e-01;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 EADPTGHSY 9
|||||||||
Qy 1 EADPTGHSY 9

RESULT 2
ID Y01727 standard; Peptide; 9 AA.
AC Y01727;
DT 25-JUN-1999 (first entry)
DE Exemplary antigenic peptide derived from MAGE-1.
KW MAGE-3; tumour associated gene; human leucocyte antigen Class II;

KW autologous CD4+ cell; MAGE-3 related disease; cancer; melanoma;
KW osteosarcoma; leukemia; carcinoma.
OS Homo sapiens.
PN WO9914326-A1.
PD 25-MAR-1999.
PF 04-SEP-1998; U18601.
PR 12-SEP-1997; US-928615.
PA (LUDW-) LUDWIG INST CANCER RES.
PA (UYVR-) UNIV VRIJE BRUSSEL.
PI Boon-Falleur T, Chaux P, Corthals J, Heirman C,
PI Luiten R, Stroobant V, Thielemans K, Van Der Bruggen P;
DR WPI; 99-244031/20.
PT Isolated peptides that bind to human leucocyte antigen class II
PT molecules
PS Disclosure; Page 27; 88pp; English.
CC The present sequence represents an exemplary tumour associated peptide
CC antigen. The specification describes a MAGE-3 tumour associated gene.
CC Peptides (Y01721-25) that bind human leucocyte antigen (HLA) Class II
CC molecules can be derived from the MAGE-3 protein. These peptides and
CC autologous CD4+ cells that bind to a complex of MAGE-3 peptide
CC and HLA Class II, are used to treat MAGE-3 related diseases,
CC particularly cancers (e.g. melanoma, osteosarcoma, leukemia and
CC various forms of carcinoma). The peptides are also used to produce
CC specific antibodies. Detection of of the peptides, e.g. in binding
CC assays, particularly with antibodies, is used for diagnosis of such
CC diseases.
SQ Sequence 9 AA;

Query Match 100.0%; Score 61; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.32e-01;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 EADPTGHSY 9
|||||||||
Qy 1 EADPTGHSY 9

RESULT 3
ID Y10633 standard; Peptide; 9 AA.
AC Y10633;
DT 12-MAY-1999 (first entry)
DE Peptide antigen SEQ ID NO:563.
KW Cytotoxic T-lymphocyte response; CTL; antigen; lymphatic system;
KW immunisation; tumour; infectious disease; immunotherapy; cancer;
KW malignant melanoma; viral disease; hepatitis; AIDS.
OS Synthetic.
OS Homo sapiens.
PN WO9902183-A2.
PD 21-JAN-1999.
PF 10-JUL-1998; U14289.
PR 10-DEC-1997; US-988320.
PR 10-JUL-1997; CA-209815.
PA (CTLI-) CTL IMMUNOTHERAPIES CORP.
PI Kuendig TM, Simard JJL;
DR WPI; 99-120514/10.
PT Inducing a cytotoxic T lymphocyte response - by maintaining a level
PT of antigen in the lymphatic system of a mammal so as to provide a
PT sustained CTL response, used to treat, e.g. AIDS
PS Disclosure; Page 52; 199pp; English.
CC The present invention describes a method of inducing and/or sustaining
CC an immunological cytotoxic T lymphocyte (CTL) response in a mammal. The
CC method comprises: (a) delivering an antigen to the mammal at a level to
CC induce an immunological CTL response in the mammal; and (b) maintaining
CC the level of the antigen in the mammal's lymphatic system to maintain
CC the immunologic CTL response. The method can be used for the delivery of
CC e.g. a differentiation antigen, a tumour-specific multilineage antigen,
CC an embryonic antigen, an oncogene antigen, a mutated tumour-suppressor
CC gene antigen, or a viral antigen. They can be used for the treatment of
CC disease such as cancer, e.g. malignant melanoma or infectious disease,
CC e.g. viral disease such as hepatitis or AIDS. Sustained antigen delivery
CC to the lymphatic system provides for potent CTL stimulation that takes
CC place in the milieu of the lymphoid organ, and it sustains stimulation
CC that is necessary to keep CTL active, cytotoxic and recirculating
CC through the body. Y10071 to Y10639 represent examples of peptide
CC antigens given in the present invention.
SQ Sequence 9 AA;

Query Match 100.0%; Score 61; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.32e-01;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 EADPTGHSY 9
|||||||||
Qy 1 EADPTGHSY 9

RESULT 4
ID Y00685 standard; peptide; 9 AA.
AC Y00685;
DT 12-MAY-1999 (first entry)
DE Tumour antigen booster peptide MAGE-1 HLA-A1.
KW Tumour antigen; booster peptide; immune response modulation; allergy;
KW immune response enhancer; tumour cell; tumour rejection antigen;
KW leukocyte antigen-presenting molecule; autoimmune disease;
KW allograft rejection.
OS Homo sapiens.
PN WO9858956-A2.
PD 30-DEC-1998.
PF 19-JUN-1998; U12894.
PR 23-JUN-1997; US-880979.
PA (LUDW-) LUDWIG INST CANCER RES.
PI Boon-Falleur T, Uyttenhove C, Warnier G;
DR WPI; 99-105612/09.
PT Immunization methods using viruses expressing antigen for priming
PT and booster immunizations - useful for modulating immune responses
PT against antigen, e.g. enhancing immune response against tumour cells
PT expressing tumour rejection antigens
PS Claim 3; Page 9; 33pp; English.
CC This sequence represents a tumour antigen booster peptide that can be
CC used in the method of the invention. The method is for for modulating an
CC immune response in a mammal against an antigen, and comprises:
CC (A) inducing an immune response by: (i) administering a virus containing
CC a nucleic acid molecule encoding the antigen or its precursor to generate
CC an immune response; and (ii) administering at least one booster dose
CC comprising a peptide including the antigen, in an adjuvant, in a combined
CC amount effective to enhance the initial immune response; or
CC (B) reducing an immune response as defined for (A) but using a
CC non-adjuvant with the peptide which includes the antigen, in an amount
CC effective to reduce the initial immune response. Method (A) is used to
CC enhance the immune response against tumour cells expressing tumour
CC rejection antigens, and against pathogens in subjects having human
CC leukocyte antigen-presenting molecules. Method (B) is used to reduce the
CC immune response in allergy, autoimmune disease, and allograft rejection.
CC Method (A) provides an immunisation method which, unlike prior art, is
CC not limited by the host immune response against viral vectors.
SQ Sequence 9 AA;

Query Match 100.0%; Score 61; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.32e-01;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 EADPTGHSY 9
|||||||||
Qy 1 EADPTGHSY 9

RESULT 5
ID Y10424 standard; Peptide; 9 AA.
AC Y10424;
DT 12-MAY-1999 (first entry)
DE HLA Class I motif peptide SEQ ID NO:354.
KW Cytotoxic T-lymphocyte response; CTL; antigen; lymphatic system;
KW immunisation; tumour; infectious disease; immunotherapy; cancer;
KW malignant melanoma; viral disease; hepatitis; AIDS.
OS Synthetic.
OS Homo sapiens.

PN WO9902183-A2.
PD 21-JAN-1999.
PF 10-JUL-1998; U14289.
PR 10-DEC-1997; US-988320.
PR 10-JUL-1997; CA-209815.
PA (CTLI-) CTL IMMUNOTHERAPIES CORP.
PI Kuendig TM, Simard JJL;
DR WPI; 99-120514/10.
PT Inducing a cytotoxic T lymphocyte response - by maintaining a level
PT of antigen in the lymphatic system of a mammal so as to provide a
PT sustained CTL response, used to treat, e.g. AIDS
PS Disclosure; Page 39; 199pp; English.
CC The present invention describes a method of inducing and/or sustaining
CC an immunological cytotoxic T lymphocyte (CTL) response in a mammal. The
CC method comprises: (a) delivering an antigen to the mammal at a level to
CC induce an immunological CTL response in the mammal; and (b) maintaining
CC the level of the antigen in the mammal's lymphatic system to maintain
CC the immunologic CTL response. The method can be used for the delivery of
CC e.g. a differentiation antigen, a tumour-specific multilineage antigen,
CC an embryonic antigen, an oncogene antigen, a mutated tumour-suppressor
CC gene antigen, or a viral antigen. They can be used for the treatment of
CC disease such as cancer, e.g. malignant melanoma or infectious disease,
CC e.g. viral disease such as hepatitis or AIDS. Sustained antigen delivery
CC to the lymphatic system provides for potent CTL stimulation that takes
CC place in the milieu of the lymphoid organ, and it sustains stimulation
CC that is necessary to keep CTL active, cytotoxic and recirculating
CC through the body. Y10071 to Y10639 represent examples of peptide
CC antigens given in the present invention.
SQ Sequence 9 AA;

Query Match 100.0%; Score 61; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.32e-01;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 EADPTGHSY 9
||||||||| |
Qy 1 EADPTGHSY 9

RESULT 6
ID Y10623 standard; Peptide; 9 AA.
AC Y10623;
DT 12-MAY-1999 (first entry)
DE Peptide antigen SEQ ID NO:553.
KW Cytotoxic T-lymphocyte response; CTL; antigen; lymphatic system;
KW immunisation; tumour; infectious disease; immunotherapy; cancer;
KW malignant melanoma; viral disease; hepatitis; AIDS.
OS Synthetic.
OS Homo sapiens.
PN WO9902183-A2.
PD 21-JAN-1999.
PF 10-JUL-1998; U14289.
PR 10-DEC-1997; US-988320.
PR 10-JUL-1997; CA-209815.
PA (CTLI-) CTL IMMUNOTHERAPIES CORP.
PI Kuendig TM, Simard JJL;
DR WPI; 99-120514/10.
PT Inducing a cytotoxic T lymphocyte response - by maintaining a level
PT of antigen in the lymphatic system of a mammal so as to provide a
PT sustained CTL response, used to treat, e.g. AIDS
PS Disclosure; Page 51; 199pp; English.
CC The present invention describes a method of inducing and/or sustaining
CC an immunological cytotoxic T lymphocyte (CTL) response in a mammal. The
CC method comprises: (a) delivering an antigen to the mammal at a level to
CC induce an immunological CTL response in the mammal; and (b) maintaining
CC the level of the antigen in the mammal's lymphatic system to maintain
CC the immunologic CTL response. The method can be used for the delivery of
CC e.g. a differentiation antigen, a tumour-specific multilineage antigen,
CC an embryonic antigen, an oncogene antigen, a mutated tumour-suppressor
CC gene antigen, or a viral antigen. They can be used for the treatment of
CC disease such as cancer, e.g. malignant melanoma or infectious disease,
CC e.g. viral disease such as hepatitis or AIDS. Sustained antigen delivery
CC to the lymphatic system provides for potent CTL stimulation that takes
CC place in the milieu of the lymphoid organ, and it sustains stimulation
CC that is necessary to keep CTL active, cytotoxic and recirculating
CC through the body. Y10071 to Y10639 represent examples of peptide
CC antigens given in the present invention.
SQ Sequence 9 AA;

Query Match 100.0%; Score 61; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.32e-01;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 EADPTGHSY 9
||||||||| |
Qy 1 EADPTGHSY 9

RESULT 7
ID W54622 standard; peptide; 9 AA.
AC W54622;
DT 25-SEP-1998 (first entry)
DE Peptide from Mage-1 161-169.
KW Mannose; antigen; antigen-presenting cell; mannosylated peptide; T cell;
KW vaccine; treatment.
OS Synthetic.
PN WO9813378-A1.
PD 02-APR-1998.
PF 25-SEP-1997; NL0536.
PR 26-SEP-1996; EP-202701.
PA (UYLE-) RIJKSUNIV LEIDEN.
PI Drijfhout JW, Koning F;
DR WPI; 98-230631/20.
PT Increasing uptake and presentation of antigen(s) - by adding mannose
PT residue(s) to antigen for increasing T cell response, useful in,
PT e.g. vaccines against viral infection(s)
PS Disclosure; Page 28; 47pp; English.
CC The peptides W54559-W54809 are examples of peptides to which at least 1
CC (preferably 2) mannose can be attached to increase their uptake as
CC antigens by antigen-presenting cells. Uptake of agonist mannosylated
CC peptides will increase the T cell response, whereas uptake of antagonist
CC peptides blocks the T cell response. Blocking binding of immunogenic
CC autoantigens can be used in treatment of type I diabetes, rheumatoid
CC arthritis, graft rejection etc., also to induce T-cell non-
CC responsiveness. Vaccines containing mannosylated antigen are used to
CC prevent or treat infections by, e.g. bacteria, viruses, fungi, helminths
CC and parasites.
SQ Sequence 9 AA;

Query Match 100.0%; Score 61; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.32e-01;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 EADPTGHSY 9
||||||||| |
Qy 1 EADPTGHSY 9

RESULT 8
ID R50281 standard; Protein; 9 AA.
AC R50281;
DT 26-SEP-1994 (first entry)
DE MAGE-1 nonapeptide.
KW MAGE; nonapeptide; cancer; melanoma; breast cancer; HLA;
KW histocompatability; human leucocyte antigen; probe; treatment;
KW therapy; vaccine.
OS Synthetic.
PN WO9405304-A.
PD 17-MAR-1994.
PF 30-AUG-1993; U08157.
PR 31-AUG-1992; US-938334.
PR 26-MAR-1993; US-037230.
PR 07-JUN-1993; US-073103.
PA (LUDW-) LUDWIG INST CANCER RES.
PI Boon-falleur T, De Plaen E, Lurquin C, Traversari C;
PI Van Derbruggen P;

```
DR   WPI; 94-100844/12.
DR   N-PSDB; Q44751.
PT   New nona:peptide derived from tumour rejection antigen precursor
PT   - presented by HLA-A1 cancer cells, for use in diagnosis or
PT   therapy of esp. melanoma and breast cancer.
PS   Disclosure; Page 19; 33pp; English.
CC   An isolated nonapeptide having the amino acid sequence Glu-Val-Asp-
CC   Pro-Ile-Gly-His-Leu-Tyr is derived from the tumour rejection antigen
CC   precursor encoded by the MAGE-3 gene and presented by HLA-A1.  The
CC   nonapeptide can be used in a vaccine to treat a cancerous condition
CC   involving HLA-A1 subtype cancerous cells.  The nucleic acid encoding
CC   the nonapeptide can be used as a probe to identify tumour cells.
CC   This sequence is homologous to the peptide described and is encoded
CC   by the MAGE-1 gene.
SQ   Sequence   9 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   9
ID   R83932 standard; peptide; 9 AA.
AC   R83932;
DT   05-JUN-1996  (first entry)
DE   MHC class I restricted antigenic peptide #2.
KW   MHC class I; antigen; MAGE; melanoma; breast cancer; bladder cancer;
KW   Titermax; cytotoxic T-lymphocyte; tumour; pathogenic disease; bacteria;
KW   parasite; human; animal.
OS   Synthetic.
PN   WO9528958-A1.
PD   02-NOV-1995.
PF   21-APR-1995; U04975.
PR   22-APR-1994; US-233496.
PA   (SLOK ) SLOAN KETTERING INST CANCER RES.
PI   Dyall R,  Nikolic-Zugic J;
DR   WPI; 95-382848/49.
PT   Cytotoxic T-cell induction by MHC class I-restricted peptide in
PT   adjuvant - useful for treating tumours and bacterial or parasitic
PT   pathogenic diseases
PS   Claim 11; Page 38; 50pp; English.
CC   The sequences given in R83931-49 are MHC class I restricted 8-12
CC   amino acid antigenic peptides.  This peptide is derived from MAGE
CC   and is present in melanoma, breast and bladder cancer.  These
CC   peptides may be administered to a subject in combination with a
CC   suitable adjuvant, pref. Titermax (RTM), to induce cytotoxic T-
CC   lymphocytes.  This method may be used in the treatment of a tumour
CC   or a pathogenic disease, esp. diseases of bacterial or parasitic
CC   origin, in humans and animals, e.g monkeys, dogs cows, horses, etc.
SQ   Sequence   9 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   10
ID   R29769 standard; Peptide; 9 AA.
AC   R29769;
DT   22-APR-1993  (first entry)
DE   Antigen E peptide.
KW   Antigen; tumorigenic cell; A+ B+; T-cell; response; syngeneic;
KW   animal; mouse; tumour rejection antigen precusor; TRAP; P1A.
OS   Homo sapiens.
PN   WO9220356-A.
PD   26-NOV-1992.
PF   22-MAY-1992; U04354.
PR   23-MAY-1991; US-705702.
PR   09-JUL-1991; US-728838.
PR   23-SEP-1991; US-764364.
PR   12-DEC-1991; US-807043.
PA   (LUDW-) LUDWIG INST CANCER RES.
PI   Boon T, Chomez P, De Plaen E, Lurquin C, Traversari C;
PI   Van Den Eynde B, Van Der Bruggen P, Van Pel A;
DR   WPI; 92-415460/50.
PT   Nucleic acid mol. encoding a human tumour rejection antigen
PT   precursor - useful as an immunostimulant in a vaccine for
PT   treating and preventing cancers, also useful in diagnosis
PS   Disclosure; Page 97; 142pp; English.
CC   This sequence represents the sequence of the antigen E.  Antigens such
CC   as this one cause a T-cell response to be elicited which transplanted
CC   into a syngeneic animal, usually a mouse.  This antigen is derived from
CC   thecell line MEL3.1.  See also Q32351.
SQ   Sequence   9 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   11
ID   W68371 standard; peptide; 9 AA.
AC   W68371;
DT   14-OCT-1998  (first entry)
DE   Human MAGE-1 peptide binds HLA-A1.
KW   Antigen; major histocompatibility complex; MHC; lymphocyte; detection;
KW   immobilisation; cytotoxic T-cell; tumour; leukaemia; lymphoma;
KW   viral infection.
OS   Synthetic.
OS   Homo sapiens.
PN   WO9744667-A2.
PD   27-NOV-1997.
PF   21-MAY-1997; F00892.
PR   21-MAY-1996; US-651925.
PA   (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
PA   (INSP ) INST PASTEUR.
PI   Abastado J, Kourilsky P, Langlade-Demoyen P, Lone Y;
DR   WPI; 98-018653/02.
PT   Detection, purification and elimination of antigen-specific
PT   lymphocytes - for producing cytotoxic T cells for immuno-therapy of
PT   cancers and viral infection
PS   Disclosure; Page 30; 222pp; French.
CC   Peptides W68301-W68384 are examples of antigens (Ag) which can be loaded
CC   onto recombinantly produced major histocompatibility complex (MHC)
CC   molecules in a method of detecting antigen-specific lymphocytes.  The
CC   MHC-antigen complex is then immobilised on a solid support and a sample
CC   containing cells recognising the MHC-Ag complex may be isolated.  This
CC   peptide is derived from the human MAGE-1 protein and binds the human
CC   leukocyte antigen A1 (HLA-A1).  A similar method is used to isolate,
CC   purify or eliminate Ag-specific T-cells or to produce Ag-specific
CC   cytotoxic T-cells (CTC).  The method is also used to detect and quantify
CC   tumour-specific T-cells and to generate CTC for specific killing of
CC   tumour cells (solid tumours, leukaemia or lymphoma) by injection into
CC   a human or animal, but also for treating viral infections.
SQ   Sequence   9 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9
```

```
RESULT   12
ID   W77125 standard; peptide; 9 AA.
AC   W77125;
DT   16-NOV-1998  (first entry)
DE   gp75/TRP-1 synthetic peptide epitope 1.
KW   Tyrosinase; tyrosinase cytotoxic lymphocyte response;
KW   cytotoxic T lymphocyte; cysteine-depleted; melanoma.
OS   Synthetic.
PN   WO9833810-A2.
PD   06-AUG-1998.
PF   29-JAN-1998; U01592.
PR   30-JAN-1997; US-037781.
PA   (UYVI-) UNIV VIRGINIA PATENT FOUND.
PI   Engelhard VH, Hunt DF, Kittlesen D, Slingluff CL;
DR   WPI; 98-437388/37.
PT   Disease specific immunogen  - comprises disease specific cytotoxic T
PT   lymphocyte epitope used to elicit melanoma specific CTL response
PS   Disclosure; Page 27; 93pp; English.
CC   The peptide epitope W77119-W77138 were created for human tumour-specific
CC   cytotoxic T lymphocyte response.  These peptides are are cysteine-
CC   depleted mutants of a native disease-specific CTL epitope.  The cysteine-
CC   depleted CTL epitopes elicit a stronger or more specific CTL response
CC   than the native epitope.  The epitopes can be used in a disease-specific
CC   immunogen to protect a mammal against disease in particular melanomas.
CC   The peptides may also be used to screen a sample for the presence of
CC   an antigen with the same epitope, or with a different cross-reactive
CC   epitope.
SQ   Sequence   9 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   13
ID   W75734 standard; peptide; 9 AA.
AC   W75734;
DT   19-NOV-1998  (first entry)
DE   Peptidase-resistant peptide 2.
KW   Tumour antigen MZ2-E; T-cell; immunotherapy; cytolytic T-cell; CTL;
KW   therapeutic agent; peptidase; MZ2-E antigen peptide analogue; HLA;
KW   human leucocyte antigen; MHC; lysis; vaccine.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   Misc_difference 2
FT                   /note= "D-form residue"
FT   Misc_difference 8
FT                   /note= "D-form residue"
PN   WO9833511-A1.
PD   06-AUG-1998.
PF   19-NOV-1997; U21296.
PR   05-FEB-1997; US-795733.
PA   (CNRS ) CENT NAT RECH SCI.
PA   (LUDW-) LUDWIG INST CANCER RES.
PI   Ayyoub M, Gairin JE, Mazarguil H, Monsarrat B, Van Den Eynde B;
DR   WPI; 98-437166/37.
PT   Peptidase-resistant peptide(s) that bind to HLA molecules and
PT   related antibodies - particularly for treatment of cancer by
PT   inducing proliferation of cytotoxic T cells
PS   Claim 20; Page 20; 32pp; English.
CC   Sequences W75733-W75736 are peptidase-resistant peptides which are
CC   analogues of the tumour antigen MZ2-E.  This antigen is a potential
CC   target for T-cell based immunotherapy and can also be used to stimulate
CC   the antigen-specific CTL, however its use as a therapeutic agent is
CC   limited due to its degradation by peptidase.  The MZ2-E antigen peptide
CC   analogues were modified at both peptidase sensitive portions, and were
CC   all shown to exhibit a longer half-life relative to peptidase degradation
CC   as well as the ability to bind a human leucocyte antigen (HLA).  The
CC   specific peptides W75733 and W75735 were established to have a comparable
CC   affinity for the MHC as the tumour antigen, and W75735 was found to be
CC   the ideal peptide analog to use due to it also being able to sensitise
CC   the target cells to lysis by effector molecules at similar concentrations
CC   to those of the antigen MZ2-E.  These peptide analogues can be used in
CC   vaccines to induce an immune response for treating conditions in which
CC   abnormal HLA/peptide complexes are present on the surface of cells.
SQ   Sequence   9 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   14
ID   W00897 standard; Peptide; 9 AA.
AC   W00897;
DT   23-MAY-1997  (first entry)
DE   Human melanoma MAGE1 tumour associated antigen p161-169.
KW   Adeno-associated virus; vector; liposome; transfection;
KW   dendritic cell; melanoma; MAGE1; adoptive immunotherapy;
KW   tumour associated antigen.
OS   Homo sapiens.
PN   WO9703703-A1.
PD   06-FEB-1997.
PF   19-JUL-1996; U12012.
PR   21-JUL-1995; US-001312.
PR   01-NOV-1995; US-007184.
PR   01-DEC-1995; US-566286.
PA   (RHON ) RHONE POULENC RORER PHARM INC.
PI   Lebkowski JS,  Philip R;
DR   WPI; 97-145208/13.
PT   Adeno-associated virus:liposome complexes for transfecting dendritic
PT   cells - for inducing immune response, useful for treating e.g.
PT   neoplasia or infections
PS   Example 5; Page 58; 134pp; English.
CC   Tumour associated antigens (W13660-61, W00878-903) can be loaded
CC   into dendritic cells and used to induce antitumour immunity.
CC   Alternatively, the dendritic cells are transfected with adeno
CC   associated virus plasmid DNA (which includes DNA encoding the
CC   tumour associated antigen) complexed with cationic liposomes.  The
CC   amtigen loaded or transfected dendritic cells can be used to
CC   generate tumour antigen-specific cytotoxic T lymphocytes for use in
CC   adoptive immunotherapy in a patient having the corresponding
CC   tumour.  A suitable antigen comprises amino acids 161-169 (W00897)
CC   of human melanoma MAGE1.
SQ   Sequence   9 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   15
ID   W75736 standard; peptide; 9 AA.
AC   W75736;
DT   19-NOV-1998  (first entry)
DE   Peptidase-resistant peptide 4.
KW   Tumour antigen MZ2-E; T-cell; immunotherapy; cytolytic T-cell; CTL;
KW   therapeutic agent; peptidase; MZ2-E antigen peptide analogue; HLA;
KW   human leucocyte antigen; MHC; lysis; vaccine.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   Modified_site   2
FT                   /note= "N-Methyl-Alanine"
```

```
FT   Modified_site   8
FT                   /note= "N-Methyl-Serine"
PN   WO9833511-A1.
PD   06-AUG-1998.
PF   19-NOV-1997; U21296.
PR   05-FEB-1997; US-795733.
PA   (CNRS ) CENT NAT RECH SCI.
PA   (LUDW-) LUDWIG INST CANCER RES.
PI   Ayyoub M, Gairin JE, Mazarguil H, Monsarrat B, Van Den Eynde B;
DR   WPI; 98-437166/37.
PT   Peptidase-resistant peptide(s) that bind to HLA molecules and
PT   related antibodies - particularly for treatment of cancer by
PT   inducing proliferation of cytotoxic T cells
PS   Claim 20; Page 20; 32pp; English.
CC   Sequences W75733-W75736 are peptidase-resistant peptides which are
CC   analogues of the tumour antigen MZ2-E.  This antigen is a potential
CC   target for T-cell based immunotherapy and can also be used to stimulate
CC   the antigen-specific CTL, however its use as a therapeutic agent is
CC   limited due to its degradation by peptidase.  The MZ2-E antigen peptide
CC   analogues were modified at both peptidase sensitive portions, and were
CC   all shown to exhibit a longer half-life relative to peptidase degradation
CC   as well as the ability to bind a human leucocyte antigen (HLA).  The
CC   specific peptides W75733 and W75735 were established to have a comparable
CC   affinity for the MHC as the tumour antigen, and W75735 was found to be
CC   the ideal peptide analog to use due to it also being able to sensitise
CC   the target cells to lysis by effector molecules at similar concentrations
CC   to those of the antigen MZ2-E.  These peptide analogues can be used in
CC   vaccines to induce an immune response for treating conditions in which
CC   abnormal HLA/peptide complexes are present on the surface of cells.
SQ   Sequence   9 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   16
ID   R82988 standard; Peptide; 9 AA.
AC   R82988;
DT   26-FEB-1996  (first entry)
DE   P815 antigenic peptide.
KW   P815 antigen; P1A antigen; cancer; vaccine.
OS   Synthetic.
PN   WO9523874-A1.
PD   08-SEP-1995.
PF   23-FEB-1995; U02203.
PR   01-MAR-1994; US-204727.
PR   10-MAR-1994; US-209172.
PR   01-SEP-1994; US-299849.
PR   30-NOV-1994; US-346774.
PA   (LUDW-) LUDWIG INST CANCER RES.
PI   Boon-Falleur T,  Brasseur F,  Chomez P,  De Plaen E;
PI   De Smet C,  Gaugler B, Lethe B,  Marchand M,  Patard J;
PI   Szikora J,  Van Den Eynde B,  Van Derbruggen P,  Weynants P;
DR   WPI; 95-320586/41.
PT   Determn. of cancerous condition(s) - using a nucleic acid as a
PT   primer to determine expression of a MAGE tumour rejection antigen
PT   precursor
PS   Example 13; Page 22; 121pp; English.
CC   Using the sequence of the P815A antigen precursor gene P1A
CC   (T01176), an antigenic peptide (R82988) which was A+B+ (i.e.
CC   characteristic of cells which express both A and B antigens) was
CC   produced.  The peptide lysed PO.HTR cells in the presence of
CC   cytolytic T lymphocyte cell lines, and may be useful as a vaccine
CC   component.
SQ   Sequence   9 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   17
ID   R90692 standard; peptide; 9 AA.
AC   R90692;
DT   31-JUL-1996  (first entry)
DE   Human leukocyte antigen (HLA-A1) presented peptide MZ2-E.
KW   Human leukocyte antigen; HLA-A1; MAGE-1 derived;
KW   blood mononuclear cell; BMC; CD8-beta+ cell; cytolytic T cell;
KW   CTL cell; treatment; tumour cell; diagnosis; assay;
KW   presented peptide.
OS   Synthetic.
PN   WO9535500-A1.
PD   28-DEC-1995.
PF   14-JUN-1995; U07559.
PR   17-JUN-1994; US-261541.
PA   (LUDW-) LUDWIG INST CANCER RES.
PI   Boon-Falleur T,  Coulie P,  Van Der Bruggen P;
DR   WPI; 96-058510/06.
PT   Prodn. of specific cytolytic T cell sub-populations - by contacting
PT   blood mononuclear cells with specific peptide(s) and a population of
PT   CD8-beta(+) cells
PS   Claim 5; Page 19; 25pp; English.
CC   The present peptide is the human leukocyte antigen (HLA-A1), MAGE-1
CC   derived presented peptide, MZ2-E. By contacting a sample of blood
CC   mononuclear cells (BMC) with the peptide (which binds directly to
CC   HLA-A1 mols. on the surface of the BMC) and CD8-beta+ cells (which
CC   stimulate peptide/HLA-A1 complex specific CD8-beta+ cells), a
CC   peptide/HLA-A1 complex specific cytolytic T (CTL) cell
CC   subpopulation can be obtd. . The CTL cells obtd. can be
CC   administered to a patient to treat tumour cell related conditions,
CC   and can be used in diagnostic methods, e.g. in assays for the
CC   peptide/HLA-A1 complex.
SQ   Sequence   9 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   18
ID   W56729 standard; peptide; 9 AA.
AC   W56729;
DT   31-JUL-1998  (first entry)
DE   MAGE-1 antigenic partial peptide sequence (residues 161-169).
KW   MAGE; replication defective; adenovirus; tumour; antigen; cancer;
KW   immunotherapy; tumour rejection antigen precursor; TRAP; CTL;
KW   human leukocyte antigen; HLA; cytolytic T lymphocyte.
OS   Synthetic.
PN   WO9815638-A2.
PD   16-APR-1998.
PF   06-OCT-1997; U17948.
PR   06-OCT-1996; US-027891.
PA   (LUDW-) LUDWIG INST CANCER RES.
PI   Cerrottini J, Jongeneel CV, Reed DS, Rimoldi D,
PI   Romero P;
DR   WPI; 98-240824/21.
PT   New replication-defective adenoviruses - comprise insert encoding
PT   tumour rejection antigen precursor(s), useful for, e.g. cancer
PT   immuno-therapy
PS   Examples; Page 42; 56pp; English.
CC   This is a partial sequence of the MAGE-1 antigenic peptide used in the
CC   methods of the invention. The specification provides a new nucleic acid
CC   molecule comprising a replication-defective adenovirus genome containing
```

```
CC   an insert encoding a tumour rejection antigen precursor (TRAP). The
CC   replication-defective adenovirus genome is useful as a vector for
CC   introducing a TRAP molecule into mammalian (especially human) cells. The
CC   recombinant adenovirus is preferably targetted to tumour cells, e.g. by
CC   binding a ligand to the virus coat. The TRA peptides which are generated
CC   from the expressed TRAP are presented by human leukocyte antigen (HLA)
CC   molecules and as a result cytolytic T lymphocyte (CTL) production is
CC   increased (claimed). The CTL's then kill the TRAP-expressing tumour
CC   cells. Also, cells transfected by the recombinant adenovirus can be used
CC   for assessing the processing of TRAPs, including post-translational
CC   modifications. The adenovirus (genome) can be administered by injection,
CC   topical application or intracavitarily in 106-1010 pfu doses. The range
CC   of TRA peptides produced by replication-defective adenovirus means that
CC   patients with a range of HLA phenotypes can be treated. Also, host cell
CC   immune response to TRA's is enhanced, e.g. by induction of tumour-
CC   specific cytolytic T lymphocytes.
SQ   Sequence   9 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   19
ID   R75954 standard; Peptide; 9 AA.
AC   R75954;
DT   06-MAR-1996  (first entry)
DE   Melanoma antigen (MAGE-1) epitope.
KW   MAGE-3; melanoma antigen; vaccine; immune response; immunogenic peptide;
KW   cytotoxic T lymphocyte response; CTL; melanoma; breast cancer; antibody.
OS   Homo sapiens.
PN   WO9519783-A1.
PD   27-JUL-1995.
PF   25-JAN-1995; U01000.
PR   25-JAN-1994; US-186266.
PA   (CYTE-) CYTEL CORP.
PI   Celis E,  Grey HM,  Kubo RT,  Sette A;
DR   WPI; 95-269270/35.
PT   Immunogenic peptide(s) that induce immune response to cancer cells
PT   - that express a MAGE-3 protein peptide epitope used in vaccines or
PT   adoptive immuno:therapy to induce cytotoxic T lymphocytes
PS   Example; Page 33; 44pp; English.
CC   R75954 is derived from MAGE-1 protein.  It was used to show the
CC   specificity of CTL response to MAGE-3 peptides shown in R75942-53.
CC   R75942 is derived from the sequence of the melanoma antigen (MAGE-3)
CC   protein and can be used to elicit a primary cytotoxic T lymphocyte
CC   response against cells expressing MAGE-3. Synthetic peptides R75945-53
CC   can be used therapeutically to elicit CTL responses to melanoma, breast,
CC   colon, prostate, or other cells which express proteins with this epitope.
CC   The peptides have specific HLA-A1 binding capacity.
SQ   Sequence   9 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   20
ID   R99343 standard; Protein; 9 AA.
AC   R99343;
DT   22-APR-1997  (first entry)
DE   MAGE-1 nonapeptide.
KW   HLA binding peptide; cell lysis; cytolytic T cell; MAGE family; human;
KW   tumour rejection antigen precursor; TRA; MAGE-1; tumour; cancer cell;
KW   antibody; melanoma; universal effector cell; vaccine; breast cancer; CTL;
KW   therapy.
OS   Homo sapiens.
PN   WO9626214-A1.
PD   29-AUG-1996.
PF   01-FEB-1996; U01489.
PR   23-FEB-1995; US-393273.
PA   (LUDW-) LUDWIG INST CANCER RES.
PI   Boon-Falleur T,  De Plaen E,  Gaugler B,  Lurquin C;
PI   Romero P,  Traversari C, Van Den Eynde B,  Van Der Bruggen P;
DR   WPI; 96-402317/40.
DR   N-PSDB; T35408.
PT   New nona:peptide(s) that bind to HLA molecule(s) and induce lysis -
PT   by specific cytolytic T cells, for diagnosis and treatment of
PT   tumours and to expand T cells in vitro.
PS   Example 4; Fig 4; 41pp; English.
CC   R99343-R99350 represent MAGE nonapeptides, based on the tumour rejection
CC   antigen region of the full length MAGE sequences.  These peptides were
CC   used to design the nonapeptides of the invention (see R99337-R99342),
CC   which bind to a HLA molecule on a cell, and provoke lysis by cytolytic T
CC   cells (CTLs) specific for a complex of the HLA molecule and nonapeptide.
CC   The nonapeptides can be used diagnostically to identify tumours
CC   expressing a particular HLA molecule, or to identify cancer cells.  The
CC   peptides can also be used therapeutically, to induce a CTL response to
CC   tumours (where the peptides are optionally coupled to tumour-specific
CC   antibodies), or to induce a response by CTLs that are otherwise inactive.
CC   The peptide sequences may also be used to expand specific CTLs in vitro
CC   for later return to the patient, such as for treating melanoma.  Tumour
CC   cells can be identified by using DNA encoding the nonapeptides as probes.
CC   Non-human cells transformed with the HLA-A1 gene and a DNA sequence
CC   encoding one of the peptides, can be used to generate CTLs, or to detect
CC   the presence of CTLs in human samples.  The non-human transformed cells,
CC   when polytransformed, are universal effector cells, and can be used in
CC   vaccines, or for treating melanoma or breast cancer.
SQ   Sequence   9 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   21
ID   W78838 standard; peptide; 9 AA.
AC   W78838;
DT   17-NOV-1998  (first entry)
DE   MAGE-1 protein fragment 161-169.
KW   Microparticle; delivery; polymeric matrix; autoantigen; tumour antigen;
KW   class II associated peptide; pathogen; gene therapy; genetic disease;
KW   infection; downregulation; immune response.
OS   Homo sapiens.
OS   Synthetic.
PN   WO9831398-A1.
PD   23-JUL-1998.
PF   22-JAN-1998; U01499.
PR   06-JAN-1998; US-003253.
PR   22-JAN-1997; US-787547.
PA   (PANG-) PANGAEA PHARM INC.
PI   Curley JM, Hedley ML, Langer RS, Lunsford LB;
DR   WPI; 98-427556/36.
PT   New preparations of microparticles - comprising a synthetic polymer
PT   matrix and nucleic acid comprising an expression vector for use in
PT   gene therapy
PS   Disclosure; Page 10; 101pp; English.
CC   A microparticle preparation (MP) has been developed, consisting of
CC   microparticles having a diameter of less than 100 mu m. The MP comprises:
CC   (a) a polymeric matrix (PM) consisting of one or more synthetic polymers
CC   having a solubility in water of less that 1 mg/l; and (b) an expression
CC   vector selected from RNA molecules (at least 50% of which are closed
CC   circles) or circular plasmid DNA (at least 50% of which are supercoiled).
CC   Also described is a MP of at most 20 microns in diameter, comprising: (a)
```

```
CC   a PM; and (b) a NAM comprising an expression control sequence operatively
CC   linked to a coding sequence, where the coding sequence encodes an
CC   expression product selected from: (1) a polypeptide at least 7 amino
CC   acids in length, having a sequence identical to the sequence of: (i) a
CC   fragment of a naturally-occurring mammalian protein; or (ii) a fragment
CC   of a naturally-occurring protein from an infectious agent which infects
CC   a mammal; (2) a peptide having a length and sequence which permits it to
CC   bind to an MHC class I or II molecule; and (3) the polypeptide or the
CC   peptide linked to a trafficking sequence. W69763 to W69765, and W78793
CC   to W78897 are peptide fragments for use in the present invention. The
CC   MPs are highly effective vehicles for the delivery of polynucleotides
CC   into phagocytic cells. They can be used for gene therapy, e.g. for
CC   treating genetic diseases, infections or tumours or for downregulating
CC   an immune response.
SQ   Sequence   9 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches     9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   22
ID   R65112 standard; peptide; 9 AA.
AC   R65112;
DT   06-OCT-1995  (first entry)
DE   MAGE 1 immunogenic peptide 161-169.
KW   MAGE 1; immunogenic peptide 161-169; cytotoxic C cells;
KW   in vitro activation; cancer; AIDS; bacterial infections; malaria;
KW   fungal infections; tuberculosis; hepatitis.
OS   Homo sapiens.
PN   WO9504817-A.
PD   16-FEB-1995.
PF   01-AUG-1994; U08672.
PR   06-AUG-1993; US-103401.
PA   (CYTE-) CYTEL CORP.
PI   Celis E,  Kubo R,  Serra H,  Tsai V,  Wentworth P;
DR   WPI; 95-090895/12.
PT   In vitro activation of cytotoxic T cells for selected killing of
PT   target cells  - for treating e.g. cancer, AIDS, hepatitis etc.by
PT   incubating them with antigen presenting cells loaded with
PT   appropriate immunogenic peptide
PS   Example 3; Page 35; 53pp; English.
CC   R65109-R65145 are immunogenic peptides, they are used in a new
CC   method for the in vitro activation of cytotoxic T cells (CTC).
CC   This is achieved by incubating the CTCs with antigen presenting
CC   cells loaded with an appropriate immunogenic peptide (e.g. one
CC   of the above peptides). By selecting the peptides used the
CC   following diseases and infections can be treated; cancer, AIDS,
CC   hepatitis, other viral and bacterial infections, malaria and
CC   tuberculosis.
SQ   Sequence   9 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches     9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   23
ID   R63675 standard; Protein; 9 AA.
AC   R63675;
DT   22-JUN-1995  (first entry)
DE   Synthetic peptide derived from exon 3.1 of MAGE 1.
KW   Melanoma antigen-1; MAGE-1; cytolytic T cells; antigen E; exon 3.1.
OS   Synthetic.
PN   WO9423031-A.
PD   13-OCT-1994.
PF   17-MAR-1994; U02877.
PR   26-MAR-1993; US-037230.
PA   (LUDW-) LUDWIG INST CANCER RES.
PI   Boon-falleur T,  Gaugler B,  Van DEN EYNDE B,  Van DER BRUGGEN P;
DR   WPI; 94-333192/41.
PT   New tumour rejection antigen precursor MAGE3 - useful in
PT   treatment and diagnosis of cancer
PS   Example 34; Page 36; 105pp; English.
CC   R63675 is a synthetic peptide derived from exon 3.1 of melanoma
CC   antigen-1 (MAGE-1), it was used to transfer antigen-E cytolytic T
CC   lymphocyte sensitivity to normally non-sensitive cells.
SQ   Sequence   9 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches     9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   24
ID   R78824 standard; peptide; 9 AA.
AC   R78824;
DT   26-MAR-1996  (first entry)
DE   MAGE-1 cytotoxic T lymphocyte epitope.
KW   MAGE-1; cytotoxic T; CTL; epitope; helper T; HTL; lymphocyte;
KW   cell; viruses; parasites; tumours; antigens; disease prevention;
KW   treatment.
OS   Homo sapiens.
PN   WO9522317-A1.
PD   24-AUG-1995.
PF   16-FEB-1995; U02121.
PR   16-FEB-1994; US-197484.
PA   (CYTE-) CYTEL CORP.
PI   Celis E,  Chesnut RW,  Grey H,  Sette AD,  Vitiello MA;
DR   WPI; 95-302545/39.
PT   Compsn. inducing cytotoxic T lymphocyte response to pref. viral,
PT   bacterial, parasitic or tumour antigens - useful in the treatment
PT   and prevention of diseases associated with the antigen e.g.
PT   hepatitis B
PS   Disclosure; Page 17; 109pp; English.
CC   A compsn. which induces a cytotoxic T lymphocyte (CTL) response to
CC   an antigen (Ag) in a mammal comprises, a CTL Ag response inducing
CC   peptide (i.e. R78824-R78853) and a lipid conjugated helper T cell
CC   inducing peptide. The compsn. induces a CTL response to bacterial,
CC   viral or tumour Ags, and is therefore useful in the treatment and
CC   prevention of diseases associated with the Ag.
SQ   Sequence   9 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches     9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   25
ID   R65135 standard; peptide; 9 AA.
AC   R65135;
DT   09-OCT-1995  (first entry)
DE   MAGE 1 immunogenic peptide A01.
KW   MAGE 1; immunogenic peptide A01; cytotoxic C cells;
KW   in vitro activation; cancer; AIDS; bacterial infections; malaria;
KW   fungal infections; tuberculosis; hepatitis.
OS   Homo sapiens.
PN   WO9504817-A.
PD   16-FEB-1995.
PF   01-AUG-1994; U08672.
```

```
PR   06-AUG-1993; US-103401.
PA   (CYTE-) CYTEL CORP.
PI   Celis E,  Kubo R,  Serra H,  Tsai V,  Wentworth P;
DR   WPI; 95-090895/12.
PT   In vitro activation of cytotoxic T cells for selected killing of
PT   target cells  - for treating e.g. cancer, AIDS, hepatitis etc.by
PT   incubating them with antigen presenting cells loaded with
PT   appropriate immunogenic peptide
PS   Example 3; Page 38; 53pp; English.
CC   R65109-R65145 are immunogenic peptides, they are used in a new
CC   method for the in vitro activation of cytotoxic T cells (CTC).
CC   This is achieved by incubating the CTCs with antigen presenting
CC   cells loaded with an appropriate immunogenic peptide (e.g. one
CC   of the above peptides). By selecting the peptides used the
CC   following diseases and infections can be treated; cancer, AIDS,
CC   hepatitis, other viral and bacterial infections, malaria and
CC   tuberculosis.
SQ   Sequence   9 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   26
ID   R49224 standard; Protein; 9 AA.
AC   R49224;
DT   31-AUG-1994  (first entry)
DE   HLA-A1 MAGE 1 antigen peptide fragment 958.01.
KW   Immunogenic; HLA-A3.2; HLA-A1; HLA-A11; binding motif; MHC molecule;
KW   immune response; viral infection; cancer; prostate cancer; lymphoma;
KW   hepatitis; AIDS; antibody; diagnosis; melanoma antigen.
OS   Synthetic.
PN   WO9403205-A.
PD   17-FEB-1994.
PF   06-AUG-1993; U07421.
PR   07-AUG-1992; US-926666.
PR   05-MAR-1993; US-027746.
PA   (CYTE-) CYTEL CORP.
PI   Celis E,  Grey HM,  Kubo RT,  Sette A;
DR   WPI; 94-065403/08.
PT   Peptide which specifically binds selected MHC allele - used to
PT   induce an immune response for treatment or prevention of viral
PT   infection or cancer, or for diagnosis
PS   Eaxmple 16; Page 116; 150pp; English.
CC   The sequences given in R47304-33 and R49201-44 are immunogenic
CC   peptides which have a HLA-A3.2, HLA-A1 or a HLA-A11 binding motif.
CC   These peptides may be used in the composition of the invention.
CC   These peptides are capable of binding selected MHC molecules and
CC   inducing an immune response.  They can be used to treat and/or
CC   prevent viral infection and cancer, eg. prostate cancer, lymphoma,
CC   hepatitis or AIDS.  They can also be used to produce antibodies for
CC   use as diagnostic or therapeutic agents.  The peptides can also be
CC   used as diagnostic agents.
SQ   Sequence   9 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   27
ID   R47330 standard; Protein; 9 AA.
AC   R47330;
DT   31-AUG-1994  (first entry)
DE   HLA-A1 MAGE 1 antigen peptide fragment 161-169.
KW   Immunogenic; HLA-A3.2; HLA-A1; HLA-A11; binding motif; MHC molecule;
KW   immune response; viral infection; cancer; prostate cancer; lymphoma;
KW   hepatitis; AIDS; antibody; diagnosis; melanoma antigen.
OS   Synthetic.
PN   WO9403205-A.
PD   17-FEB-1994.
PF   06-AUG-1993; U07421.
PR   07-AUG-1992; US-926666.
PR   05-MAR-1993; US-027746.
PA   (CYTE-) CYTEL CORP.
PI   Celis E,  Grey HM,  Kubo RT,  Sette A;
DR   WPI; 94-065403/08.
PT   Peptide which specifically binds selected MHC allele - used to
PT   induce an immune response for treatment or prevention of viral
PT   infection or cancer, or for diagnosis
PS   Eaxmple 8; Page 52; 150pp; English.
CC   The sequences given in R47304-33 and R49201-44 are immunogenic
CC   peptides which have a HLA-A3.2, HLA-A1 or a HLA-A11 binding motif.
CC   These peptides may be used in the composition of the invention.
CC   These peptides are capable of binding selected MHC molecules and
CC   inducing an immune response.  They can be used to treat and/or
CC   prevent viral infection and cancer, eg. prostate cancer, lymphoma,
CC   hepatitis or AIDS.  They can also be used to produce antibodies for
CC   use as diagnostic or therapeutic agents.  The peptides can also be
CC   used as diagnostic agents.
SQ   Sequence   9 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   28
ID   W23038 standard; Peptide; 10 AA.
AC   W23038;
DT   25-FEB-1998  (first entry)
DE   MAGE-1/HLA-B44 tumour rejection antigen.
KW   MAGE-1; tumour rejection antigen precursor; TRAP; HLA-B44;
KW   human leukocyte antigen B44; cytotoxic T lymphocyte; cancer;
KW   melanoma; therapy; diagnosis; vaccine.
OS   Homo sapiens.
PN   WO9731017-A1.
PD   28-AUG-1997.
PF   05-FEB-1997; U01915.
PR   20-FEB-1996; US-602506.
PA   (LUDW-) LUDWIG INST CANCER RES.
PI   Boon-Falleur T,  Coulie P,  Herman J,  Luescher I;
PI   Van Der Bruggen P;
DR   WPI; 97-435086/40.
PT   Tumour rejection antigens presented by human leukocyte antigen B44
PT   molecules - useful to identify HLA-B44 positive cells for diagnosis
PT   and therapy of cellular abnormalities
PS   Claim 2; Page 49; 74pp; English.
CC   This peptide is a tumour rejection antigen presented by a HLA-B44
CC   molecule and derived from a MAGE-1 tumour rejection antigen
CC   precursor (TRAP).  Claimed tumour rejection antigens (W23038-43)
CC   are able to bind to HLA-B44 positive cells, making them useful in
CC   identifying cells which present HLA-B44 molecules on their
CC   surfaces for use in the diagnosis and therapy of cellular
CC   abnormalities.  The complex of the tumour rejection antigen and HLA
CC   molecule provokes a cytolytic T cell response.  The tumour
CC   rejection antigens, or complexes of tumour rejection antigens and
CC   HLA-B44, can be used as vaccines to treat disorders characterised
CC   by expression of the TRAP molecule such as cancer, especially
CC   melanoma.  Vaccines can also be prepared from cells which present
CC   the tumour rejection antigen/HLA complexes on their surface, such
CC   as non-proliferative cancer cells and non-proliferative
CC   transfectants.
```

```
SQ   Sequence   10 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 10;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      2 EADPTGHSY 10
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   29
ID   R80620 standard; Protein; 12 AA.
AC   R80620;
DT   28-FEB-1996  (first entry)
DE   Immunogenic peptide of tumour rejection antigen (MAGE-1).
KW   Tumour rejection antigen; MAGE-1; monoclonal antibody; MAb;
KW   diagnosis; immunoassay; cancer; immunogen; antisera.
OS   Homo sapiens.
PN   WO9520974-A1.
PD   10-AUG-1995.
PF   05-JAN-1995; U00095.
PR   01-FEB-1994; US-190411.
PA   (LUDW-) LUDWIG INST CANCER RES.
PA   (SLOK ) SLOAN KETTERING INST CANCER RES.
PA   (SLOK ) MEMORIAL SLOAN-KETTERING CANCER CENT.
PI   Boon-falleur T,  Chen Y,  Garin-chesa P,  Old LJ,  Rettig WJ;
PI   Stockert E,  Van der bruggen P;
DR   WPI; 95-283606/37.
PT   New monoclonal antibody binding specifically to MAGE-1 - useful for
PT   diagnosis and monitoring of cancer, also new hybridomas, recombinant
PT   MAGE-1 and immunogenic peptide(s)
PS   Claim 12; Page 20; 33pp; English.
CC   A monoclonal antibody directed against the tumour rejection antigen
CC   (MAGE-1) can be used to detect MAGE-1 in samples by standard
CC   immunoassay methods for diagnosis and monitoring of cancer etc. The
CC   monoclonal antibody is designated MA454 and is produced by the
CC   hybridoma deposited as ATCC HBB11540.  The monoclonal antibody is
CC   specific for MAGE-1, having no reactivity for MAGE-2 or MAGE-3.
CC   Peptide fragments of MAGE-1 (See R80618-20) may be useful as
CC   immunogens for production of the monoclonal antibody and antisera.
SQ   Sequence   12 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 12;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      4 EADPTGHSY 12
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   30
ID   R70909 standard; Protein; 309 AA.
AC   R70909;
DT   09-OCT-1995  (first entry)
DE   Human melanoma antigen MAGE-1.
KW   Human melanoma antigen; MAGE-1; vaccines; MAGE associated tumours;
KW   HLA-restricted cytotoxic T-lymphocyte activity.
OS   Homo sapiens.
PN   WO9504542-A.
PD   16-FEB-1995.
PF   02-AUG-1994; U08721.
PR   06-AUG-1993; US-103623.
PA   (CYTE-) CYTEL CORP.
PI   Fikes JD,  Livingston BD,  Sette AD,  Sidney JC;
DR   WPI; 95-090681/12.
DR   N-PSDB; Q85435.
PT   Human melanoma antigen, MAGE-1, peptide(s) - useful for
PT   stimulating immune response against melanoma
PS   Example 1; Fig 1; 59pp; English.
CC   Q85435 encodes R70909 human melanoma antigen MAGE-1, it was used
CC   to produce the C-terminal MAGE-1 peptides described in R70915 to
CC   R70969. These peptides are useful for defining epitopes that
CC   engender a HLA-restricted cytotoxic lymphocyte activity against
CC   MAGE-1 antigens. Compsns. containing these peptides can be
CC   administered, as a vaccine to patients susceptable to MAGE
CC   associated tumours, e.g. melanomas.
SQ   Sequence   309 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 309;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db    161 EADPTGHSY 169
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   31
ID   W81548 standard; Protein; 309 AA.
AC   W81548;
DT   01-MAR-1999  (first entry)
DE   Tumour rejection antigen precursor MAGE-A1.
KW   MAGE-A1; human; tumour rejection antigen precursor; TRAP;
KW   therapy; diagnosis.
OS   Homo sapiens.
PN   WO9849184-A1.
PD   05-NOV-1998.
PF   24-APR-1998; U08493.
PR   25-APR-1997; US-845528.
PA   (LUDW-) LUDWIG INST CANCER RES.
PI   Boon-Falleur T, De Smet C, Lucas S;
DR   WPI; 99-024041/02.
DR   N-PSDB; V69719.
PT   Tumour rejection antigen precursors - used for determining presence
PT   of cytolytic T cells specific for complexes of a human leukocyte
PT   antigen
PS   Disclosure; Page 50-51; 84pp; English.
CC   This is the amino acid sequence of human tumour rejection antigen
CC   precursor (TRAP) MAGE-A1.  MAGE-A1 cDNA (see V69719) shows homology
CC   to novel human MAGE-C1 cDNA (see V69720).  MAGE-C1 (see W81546) is a
CC   novel member of the MAGE family that may be recognised by cytotoxic
CC   T cells, leading to lysis of the tumour cells which express it.  It
CC   is expressed in a variety of tumours and in normal testis cells,
CC   but not by other normal cells.  The invention provides MACE-C1 and
CC   MACE-C2 nucleic acids and polypeptides, useful e.g. in a claimed
CC   method for determining the presence of cytolytic T cells specific
CC   for complexes of a human leukocyte antigen (HLA).
SQ   Sequence   309 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 309;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db    161 EADPTGHSY 169
          |||||||||
Qy      1 EADPTGHSY 9


Search completed: Wed Sep 13 07:14:24 2000
Job time : 7 secs.
```

```
********************************************************************************
                        MPSRCH (TM)
********************************************************************************
          Release 3.1A John F. Collins, Biocomputing Research Unit.
          Copyright (c) 1993-1998  University of Edinburgh, U.K.
                 Distribution rights by Oxford Molecular Ltd

MPsrch_pp   protein - protein database search, using Smith-Waterman algorithm

Run on:           Tue Sep 12 13:23:01 2000;  MasPar time 5.52 Seconds
                                             76.943 Million cell updates/sec
Tabular output not generated.

Title:            >US-08-819-669E-26
Description:      (1-9) from US08819669E.pep
Perfect Score:    61
Sequence:         1 EADPTGHSY 9

Scoring table:    PAM 150
                  Gap 15

Searched:         142080 seqs, 47172406 residues

Post-processing: Minimum Match 0%
                  Listing first 1000 summaries
                  Maximum DB seq length 9

Database:         pir64
                    1:pir1 2:pir2 3:pir3 4:pir4

Statistics:       Mean 20.662;  Variance 21.178;  scale 0.976

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                SUMMARIES
                 %
Result          Query
  No.  Score  Match Length DB  ID          Description          Pred. No.
-------------------------------------------------------------------------
    1     46   75.4      9  2  PH1299      MAGE 5 protein - huma  1.79e+00

Note: Post-processor removed 999 summaries from list due to search parameters chosen.

                               ALIGNMENTS

RESULT    1
ENTRY             PH1299     #type fragment
TITLE             MAGE 5 protein - human (fragment)
ALTERNATE_NAMES   MAGE 51 protein
ORGANISM          #formal_name Homo sapiens #common_name man
DATE              30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change
                    03-Aug-1998
ACCESSIONS        PH1299; PH1300
REFERENCE         PH1294
   #authors       Traversari, C.; van der Bruggen, P.; Luescher, I.F.; Lurquin,
                    C.; Chomez, P.; Van Pel, A.; De Plaen, E.; Amar-Costesec,
                    A.; Boon, T.
   #journal       J. Exp. Med. (1992) 176:1453-1457
   #title         A nonapeptide encoded by human gene MAGE-1 is recognized on
                    HLA-A1 by cytolytic T lymphocytes directed against tumor
                    antigen MZ2-E.
   #cross-references MUID:93018875
   #accession     PH1299
      ##molecule_type DNA
      ##residues      1-9 ##label TRA
   #accession     PH1300
      ##molecule_type DNA
      ##residues      1-9 ##label TR2
SUMMARY               #length 9  #checksum 3660

  Query Match           75.4%;  Score 46;  DB 2;  Length 9;
  Best Local Similarity  66.7%;  Pred. No. 1.79e+00;
  Matches     6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Db       1 EADPTSNTY 9
           ||||||:::|
Qy       1 EADPTGHSY 9

Search completed: Tue Sep 12 13:23:48 2000
Job time : 47 secs.
```

****

MPSRCH (TM)

****

Release 3.1A John F. Collins, Biocomputing Research Unit.

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Sep 12 13:20:57 2000; MasPar time 3.60 Seconds
77.389 Million cell updates/sec

Tabular output not generated.

Title: >US-08-819-669E-26
Description: (1-9) from US08819669E.pep
Perfect Score: 61
Sequence: 1 EADPTGHSY 9

Scoring table: PAM 150
Gap 15

Searched: 85661 seqs, 30989116 residues

Post-processing: Minimum Match 0%
Listing first 1000 summaries
Maximum DB seq length 9

Database: swiss-prot38
1:swissprot

Statistics: Mean 21.172; Variance 19.906; scale 1.064

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

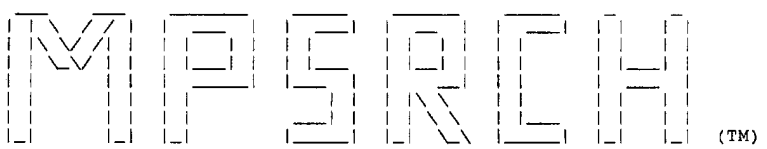
SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.

No matches found.

Search completed: Tue Sep 12 13:21:37 2000
Job time : 40 secs.

****



****

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Sep 12 13:21:54 2000; MasPar time 8.64 Seconds
72.251 Million cell updates/sec

Tabular output not generated.

Title: >US-08-819-669E-26
Description: (1-9) from US08819669E.pep
Perfect Score: 61
Sequence: 1 EADPTGHSY 9

Scoring table: PAM 150
Gap 15

Searched: 225878 seqs, 69334122 residues

Post-processing: Minimum Match 0%
Listing first 1000 summaries
Maximum DB seq length 9

Database: sptrembl12
1:sp_archea 2:sp_bacteria 3:sp_fungi 4:sp_human
5:sp_invertebrate 6:sp_mammal 7:sp_mhc 8:sp_organelle
9:sp_phage 10:sp_plant 11:sp_rodent 12:sp_unclassified
13:sp_vertebrate 14:sp_virus

Statistics: Mean 20.592; Variance 19.613; scale 1.050

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.

No matches found.

Search completed: Tue Sep 12 13:22:44 2000
Job time : 50 secs.

****

MPSRCH (TM)

****

Release 3.1A John F. Collins, Biocomputing Research Unit.

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Sep 13 06:32:58 2000; MasPar time 3.59 Seconds
59.387 Million cell updates/sec

Tabular output not generated.

Title: >US-08-819-669E-26
Description: (1-9) from US08819669E.pep
Perfect Score: 61
Sequence: 1 EADPTGHSY 9

Scoring table: PAM 150
Gap 15

Searched: 188963 seqs, 23686106 residues

Post-processing: Minimum Match 0%
Listing first 1000 summaries
Maximum DB seq length 9

Database: a-geneseq36
1:geneseqp

Statistics: Mean 15.425; Variance 35.537; scale 0.434

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	61	100.0	9	1	W98945	HLA-A1 binding peptide	2.32e-01
2	61	100.0	9	1	Y01727	Exemplary antigenic pe	2.32e-01
3	61	100.0	9	1	Y10633	Peptide antigen SEQ ID	2.32e-01
4	61	100.0	9	1	Y00685	Tumour antigen booster	2.32e-01
5	61	100.0	9	1	Y10424	HLA Class I motif pept	2.32e-01
6	61	100.0	9	1	Y10623	Peptide antigen SEQ ID	2.32e-01
7	61	100.0	9	1	W54622	Peptide from Mage-1 16	2.32e-01
8	61	100.0	9	1	R50281	MAGE-1 nonapeptide.	2.32e-01
9	61	100.0	9	1	R83932	MHC class I restricted	2.32e-01
10	61	100.0	9	1	R29769	Antigen E peptide.	2.32e-01
11	61	100.0	9	1	W68371	Human MAGE-1 peptide b	2.32e-01
12	61	100.0	9	1	W77125	gp75/TRP-1 synthetic p	2.32e-01
13	61	100.0	9	1	W75734	Peptidase-resistant pe	2.32e-01
14	61	100.0	9	1	W00897	Human melanoma MAGE1 t	2.32e-01
15	61	100.0	9	1	W75736	Peptidase-resistant pe	2.32e-01
16	61	100.0	9	1	R82988	P815 antigenic peptide	2.32e-01
17	61	100.0	9	1	R90692	Human leukocyte antige	2.32e-01
18	61	100.0	9	1	W56729	MAGE-1 antigenic parti	2.32e-01
19	61	100.0	9	1	R75954	Melanoma antigen (MAGE	2.32e-01
20	61	100.0	9	1	R99343	MAGE-1 nonapeptide.	2.32e-01
21	61	100.0	9	1	W78838	MAGE-1 protein fragmen	2.32e-01
22	61	100.0	9	1	R65112	MAGE 1 immunogenic pep	2.32e-01
23	61	100.0	9	1	R63675	Synthetic peptide deri	2.32e-01
24	61	100.0	9	1	R78824	MAGE-1 cytotoxic T lym	2.32e-01
25	61	100.0	9	1	R65135	MAGE 1 immunogenic pep	2.32e-01
26	61	100.0	9	1	R49224	HLA-A1 MAGE 1 antigen	2.32e-01
27	61	100.0	9	1	R47330	HLA-A1 MAGE 1 antigen	2.32e-01
28	59	96.7	9	1	R99342	HLA binding nonapeptid	4.62e-01
29	58	95.1	9	1	R99339	HLA binding nonapeptid	6.49e-01
30	57	93.4	9	1	W75733	Peptidase-resistant pe	9.12e-01
31	57	93.4	9	1	W75735	Peptidase-resistant pe	9.12e-01
32	55	90.2	9	1	R99337	HLA binding nonapeptid	1.79e+00
33	55	90.2	9	1	R99340	HLA binding nonapeptid	1.79e+00
34	54	88.5	9	1	R99338	HLA binding nonapeptid	2.50e+00
35	52	85.2	9	1	Y10628	Peptide antigen SEQ ID	4.86e+00
36	49	80.3	9	1	R99341	HLA binding nonapeptid	1.30e+01
37	47	77.0	9	1	Y10629	Peptide antigen SEQ ID	2.47e+01
38	46	75.4	9	1	Y10604	Peptide antigen SEQ ID	3.40e+01
39	46	75.4	9	1	R50288	MAGE-51 nonapeptide.	3.40e+01
40	46	75.4	9	1	Y10605	Peptide antigen SEQ ID	3.40e+01
41	46	75.4	9	1	R50287	MAGE-5 nonapeptide.	3.40e+01
42	46	75.4	9	1	R99349	MAGE-5/MAGE-51 nonapep	3.40e+01
43	46	75.4	9	1	R99350	MAGE-6 nonapeptide.	3.40e+01
44	46	75.4	9	1	R50289	MAGE-6 nonapeptide.	3.40e+01
45	43	70.5	9	1	Y00687	Tumour antigen booster	8.72e+01
46	43	70.5	9	1	Y01729	Exemplary antigenic pe	8.72e+01
47	43	70.5	9	1	Y10627	Peptide antigen SEQ ID	8.72e+01
48	43	70.5	9	1	R99346	MAGE-3 nonapeptide.	8.72e+01
49	43	70.5	9	1	R50284	MAGE-3 nonapeptide.	8.72e+01
50	43	70.5	9	1	W77127	MAGE-1 synthetic pepti	8.72e+01
51	43	70.5	9	1	R65118	MAGE 3 immunogenic pep	8.72e+01
52	43	70.5	9	1	W78840	MAGE-3 protein fragmen	8.72e+01
53	43	70.5	9	1	Y10427	HLA Class I motif pept	8.72e+01
54	43	70.5	9	1	W98942	MAGE-3 TRAP 168-176.	8.72e+01
55	43	70.5	9	1	Y10621	Peptide antigen SEQ ID	8.72e+01
56	43	70.5	9	1	R75942	Melanoma antigen (MAGE	8.72e+01
57	43	70.5	9	1	W68374	Human MAGE-3 peptide b	8.72e+01
58	43	70.5	9	1	R83931	MHC class I restricted	8.72e+01
59	43	70.5	9	1	W54606	Peptide 1 from Mage-3.	8.72e+01
60	43	70.5	9	1	R49222	HLA-A1 MAGE 3 antigen	8.72e+01
61	42	68.9	9	1	R99348	MAGE-41 nonapeptide.	1.19e+02
62	42	68.9	9	1	Y10603	Peptide antigen SEQ ID	1.19e+02
63	42	68.9	9	1	R50286	MAGE-41 nonapeptide.	1.19e+02
64	39	63.9	9	1	R99347	MAGE-4 nonapeptide.	2.95e+02
65	39	63.9	9	1	Y10602	Peptide antigen SEQ ID	2.95e+02
66	39	63.9	9	1	R50285	MAGE-4 nonapeptide.	2.95e+02
67	36	59.0	9	1	Y10625	Peptide antigen SEQ ID	7.11e+02
68	36	59.0	9	1	Y10626	Peptide antigen SEQ ID	7.11e+02

Note: Post-processor removed 932 summaries from list due to search parameters chosen.

ALIGNMENTS

RESULT 1
ID W98945 standard; peptide; 9 AA.
AC W98945;
DT 06-MAY-1999 (first entry)
DE HLA-A1 binding peptide derived from MAGE-1.
KW Human leukocyte antigen; HLA; HLA-A2 binding peptide; T cell;
KW cytolytic T cell; CTL.
OS Synthetic.
OS Homo sapiens.
PN WO9858951-A1.
PD 30-DEC-1998.
PF 18-JUN-1998; U12879.
PR 16-APR-1998; US-061388.
PR 23-JUN-1997; US-880963.
PA (LUDW-) LUDWIG INST CANCER RES.
PI Cerottini J, Romero P, Valmori D;
DR WPI; 99-105609/09.
PT New decamer peptides which bind to HLA molecules - useful to
PT identify HLA-A2 positive cells and provoke T cells
PS Example 7; Page 18; 45pp; English.
CC The present invention describes peptides which bind to an HLA-A2

CC molecule and have Val at the carboxy terminus, and either: (a) Ala, Tyr
CC or Phe at the amino terminus, and Ala at position 2 (P1); or (b) Glu at
CC the amino terminus, and Ala, Leu, or Met at positions 2 and 3, with the
CC proviso that Ala is not at both positions (P2). The peptides of the
CC present invention are used to identify HLA-A2 positive cells, provoke
CC T cells, and determine the presence of particular T cells including
CC cytolytic T cells (CTLs). They provide a better target than the prior
CC art CTL-stimulating peptide. The present sequence represents a peptide
CC used in an example from the present invention.
SQ Sequence 9 AA;

Query Match 100.0%; Score 61; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.32e-01;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 EADPTGHSY 9
|||||||||
Qy 1 EADPTGHSY 9

RESULT 2
ID Y01727 standard; Peptide; 9 AA.
AC Y01727;
DT 25-JUN-1999 (first entry)
DE Exemplary antigenic peptide derived from MAGE-1.
KW MAGE-3; tumour associated gene; human leucocyte antigen Class II;
KW autologous CD4+ cell; MAGE-3 related disease; cancer; melanoma;
KW osteosarcoma; leukemia; carcinoma.
OS Homo sapiens.
PN WO9914326-A1.
PD 25-MAR-1999.
PF 04-SEP-1998; U18601.
PR 12-SEP-1997; US-928615.
PA (LUDW-) LUDWIG INST CANCER RES.
PA (UYVR-) UNIV VRIJE BRUSSEL.
PI Boon-Falleur T, Chaux P, Corthals J, Heirman C,
PI Luiten R, Stroobant V, Thielemans K, Van Der Bruggen P;
DR WPI; 99-244031/20.
PT Isolated peptides that bind to human leucocyte antigen class II
PT molecules
PS Disclosure; Page 27; 88pp; English.
CC The present sequence represents an exemplary tumour associated peptide
CC antigen. The specification describes a MAGE-3 tumour associated gene.
CC Peptides (Y01721-25) that bind human leucocyte antigen (HLA) Class II
CC molecules can be derived from the MAGE-3 protein. These peptides and
CC autologous CD4+ cells that bind to a complex of MAGE-3 peptide
CC and HLA Class II, are used to treat MAGE-3 related diseases,
CC particularly cancers (e.g. melanoma, osteosarcoma, leukemia and
CC various forms of carcinoma). The peptides are also used to produce
CC specific antibodies. Detection of of the peptides, e.g. in binding
CC assays, particularly with antibodies, is used for diagnosis of such
CC diseases.
SQ Sequence 9 AA;

Query Match 100.0%; Score 61; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.32e-01;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 EADPTGHSY 9
|||||||||
Qy 1 EADPTGHSY 9

RESULT 3
ID Y10633 standard; Peptide; 9 AA.
AC Y10633;
DT 12-MAY-1999 (first entry)
DE Peptide antigen SEQ ID NO:563.
KW Cytotoxic T-lymphocyte response; CTL; antigen; lymphatic system;
KW immunisation; tumour; infectious disease; immunotherapy; cancer;
KW malignant melanoma; viral disease; hepatitis; AIDS.
OS Synthetic.
OS Homo sapiens.

PN WO9902183-A2.
PD 21-JAN-1999.
PF 10-JUL-1998; U14289.
PR 10-DEC-1997; US-988320.
PR 10-JUL-1997; CA-209815.
PA (CTLI-) CTL IMMUNOTHERAPIES CORP.
PI Kuendig TM, Simard JJL;
DR WPI; 99-120514/10.
PT Inducing a cytotoxic T lymphocyte response - by maintaining a level
PT of antigen in the lymphatic system of a mammal so as to provide a
PT sustained CTL response, used to treat, e.g. AIDS
PS Disclosure; Page 52; 199pp; English.
CC The present invention describes a method of inducing and/or sustainin
CC an immunological cytotoxic T lymphocyte (CTL) response in a mammal. T e
CC method comprises: (a) delivering an antigen to the mammal at a level o
CC induce an immunological CTL response in the mammal; and (b) maintainin
CC the level of the antigen in the mammal's lymphatic system to maintain
CC the immunologic CTL response. The method can be used for the delivery of
CC e.g. a differentiation antigen, a tumour-specific multilineage antige ,
CC an embryonic antigen, an oncogene antigen, a mutated tumour-suppresso
CC gene antigen, or a viral antigen. They can be used for the treatment f
CC disease such as cancer, e.g. malignant melanoma or infectious disease
CC e.g. viral disease such as hepatitis or AIDS. Sustained antigen deliv ry
CC to the lymphatic system provides for potent CTL stimulation that take
CC place in the milieu of the lymphoid organ, and it sustains stimulatio
CC that is necessary to keep CTL active, cytotoxic and recirculating
CC through the body. Y10071 to Y10639 represent examples of peptide
CC antigens given in the present invention.
SQ Sequence 9 AA;

Query Match 100.0%; Score 61; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.32e-01;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 EADPTGHSY 9
|||||||||
Qy 1 EADPTGHSY 9

RESULT 4
ID Y00685 standard; peptide; 9 AA.
AC Y00685;
DT 12-MAY-1999 (first entry)
DE Tumour antigen booster peptide MAGE-1 HLA-A1.
KW Tumour antigen; booster peptide; immune response modulation; allergy;
KW immune response enhancer; tumour cell; tumour rejection antigen;
KW leukocyte antigen-presenting molecule; autoimmune disease;
KW allograft rejection.
OS Homo sapiens.
PN WO9858956-A2.
PD 30-DEC-1998.
PF 19-JUN-1998; U12894.
PR 23-JUN-1997; US-880979.
PA (LUDW-) LUDWIG INST CANCER RES.
PI Boon-Falleur T, Uyttenhove C, Warnier G;
DR WPI; 99-105612/09.
PT Immunization methods using viruses expressing antigen for priming
PT and booster immunizations - useful for modulating immune responses
PT against antigen, e.g. enhancing immune response against tumour cells
PT expressing tumour rejection antigens
PS Claim 3; Page 9; 33pp; English.
CC This sequence represents a tumour antigen booster peptide that can be
CC used in the method of the invention. The method is for for modulating an
CC immune response in a mammal against an antigen, and comprises:
CC (A) inducing an immune response by: (i) administering a virus contain g
CC a nucleic acid molecule encoding the antigen or its precursor to gene ate
CC an immune response; and (ii) administering at least one booster dose
CC comprising a peptide including the antigen, in an adjuvant, in a comb ed
CC amount effective to enhance the initial immune response; or
CC (B) reducing an immune response as defined for (A) but using a
CC non-adjuvant with the peptide which includes the antigen, in an amoun
CC effective to reduce the initial immune response. Method (A) is used t
CC enhance the immune response against tumour cells expressing tumour

```
CC   rejection antigens, and against pathogens in subjects having human
CC   leukocyte antigen-presenting molecules. Method (B) is used to reduce the
CC   immune response in allergy, autoimmune disease, and allograft rejection.
CC   Method (A) provides an immunisation method which, unlike prior art, is
CC   not limited by the host immune response against viral vectors.
SQ   Sequence   9 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT    5
ID   Y10424 standard; Peptide; 9 AA.
AC   Y10424;
DT   12-MAY-1999  (first entry)
DE   HLA Class I motif peptide SEQ ID NO:354.
KW   Cytotoxic T-lymphocyte response; CTL; antigen; lymphatic system;
KW   immunisation; tumour; infectious disease; immunotherapy; cancer;
KW   malignant melanoma; viral disease; hepatitis; AIDS.
OS   Synthetic.
OS   Homo sapiens.
PN   WO9902183-A2.
PD   21-JAN-1999.
PF   10-JUL-1998; U14289.
PR   10-DEC-1997; US-988320.
PR   10-JUL-1997; CA-209815.
PA   (CTLI-) CTL IMMUNOTHERAPIES CORP.
PI   Kuendig TM, Simard JJL;
DR   WPI; 99-120514/10.
PT   Inducing a cytotoxic T lymphocyte response - by maintaining a level
PT   of antigen in the lymphatic system of a mammal so as to provide a
PT   sustained CTL response, used to treat, e.g. AIDS
PS   Disclosure; Page 39; 199pp; English.
CC   The present invention describes a method of inducing and/or sustaining
CC   an immunological cytotoxic T lymphocyte (CTL) response in a mammal. The
CC   method comprises: (a) delivering an antigen to the mammal at a level to
CC   induce an immunological CTL response in the mammal; and (b) maintaining
CC   the level of the antigen in the mammal's lymphatic system to maintain
CC   the immunologic CTL response. The method can be used for the delivery of
CC   e.g. a differentiation antigen, a tumour-specific multilineage antigen,
CC   an embryonic antigen, an oncogene antigen, a mutated tumour-suppressor
CC   gene antigen, or a viral antigen. They can be used for the treatment of
CC   disease such as cancer, e.g. malignant melanoma or infectious disease,
CC   e.g. viral disease such as hepatitis or AIDS. Sustained antigen delivery
CC   to the lymphatic system provides for potent CTL stimulation that takes
CC   place in the milieu of the lymphoid organ, and it sustains stimulation
CC   that is necessary to keep CTL active, cytotoxic and recirculating
CC   through the body. Y10071 to Y10639 represent examples of peptide
CC   antigens given in the present invention.
SQ   Sequence   9 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT    6
ID   Y10623 standard; Peptide; 9 AA.
AC   Y10623;
DT   12-MAY-1999  (first entry)
DE   Peptide antigen SEQ ID NO:553.
KW   Cytotoxic T-lymphocyte response; CTL; antigen; lymphatic system;
KW   immunisation; tumour; infectious disease; immunotherapy; cancer;
KW   malignant melanoma; viral disease; hepatitis; AIDS.
OS   Synthetic.
OS   Homo sapiens.
PN   WO9902183-A2.
PD   21-JAN-1999.
PF   10-JUL-1998; U14289.
PR   10-DEC-1997; US-988320.
PR   10-JUL-1997; CA-209815.
PA   (CTLI-) CTL IMMUNOTHERAPIES CORP.
PI   Kuendig TM, Simard JJL;
DR   WPI; 99-120514/10.
PT   Inducing a cytotoxic T lymphocyte response - by maintaining a level
PT   of antigen in the lymphatic system of a mammal so as to provide a
PT   sustained CTL response, used to treat, e.g. AIDS
PS   Disclosure; Page 51; 199pp; English.
CC   The present invention describes a method of inducing and/or sustainin
CC   an immunological cytotoxic T lymphocyte (CTL) response in a mammal. T e
CC   method comprises: (a) delivering an antigen to the mammal at a level o
CC   induce an immunological CTL response in the mammal; and (b) maintainin g
CC   the level of the antigen in the mammal's lymphatic system to maintain
CC   the immunologic CTL response. The method can be used for the delivery of
CC   e.g. a differentiation antigen, a tumour-specific multilineage antige ,
CC   an embryonic antigen, an oncogene antigen, a mutated tumour-suppresso
CC   gene antigen, or a viral antigen. They can be used for the treatment f
CC   disease such as cancer, e.g. malignant melanoma or infectious disease
CC   e.g. viral disease such as hepatitis or AIDS. Sustained antigen deliv ry
CC   to the lymphatic system provides for potent CTL stimulation that take
CC   place in the milieu of the lymphoid organ, and it sustains stimulatio
CC   that is necessary to keep CTL active, cytotoxic and recirculating
CC   through the body. Y10071 to Y10639 represent examples of peptide
CC   antigens given in the present invention.
SQ   Sequence   9 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT    7
ID   W54622 standard; peptide; 9 AA.
AC   W54622;
DT   25-SEP-1998  (first entry)
DE   Peptide from Mage-1 161-169.
KW   Mannose; antigen; antigen-presenting cell; mannosylated peptide; T ce l;
KW   vaccine; treatment.
OS   Synthetic.
PN   WO9813378-A1.
PD   02-APR-1998.
PF   25-SEP-1997; NL0536.
PR   26-SEP-1996; EP-202701.
PA   (UYLE-) RIJKSUNIV LEIDEN.
PI   Drijfhout JW, Koning F;
DR   WPI; 98-230631/20.
PT   Increasing uptake and presentation of antigen(s) - by adding mannose
PT   residue(s) to antigen for increasing T cell response, useful in,
PT   e.g. vaccines against viral infection(s)
PS   Disclosure; Page 28; 47pp; English.
CC   The peptides W54559-W54809 are examples of peptides to which at least
CC   (preferably 2) mannose can be attached to increase their uptake as
CC   antigens by antigen-presenting cells.  Uptake of agonist mannosylated
CC   peptides will increase the T cell response, whereas uptake of antagon st
CC   peptides blocks the T cell response.  Blocking binding of immunogenic
CC   autoantigens can be used in treatment of type I diabetes, rheumatoid
CC   arthritis, graft rejection etc., also to induce T-cell non-
CC   responsiveness.  Vaccines containing mannosylated antigen are used to
CC   prevent or treat infections by, e.g. bacteria, viruses, fungi, helmin hs
CC   and parasites.
SQ   Sequence   9 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
```

```
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT    8
ID   R50281 standard; Protein; 9 AA.
AC   R50281;
DT   26-SEP-1994  (first entry)
DE   MAGE-1 nonapeptide.
KW   MAGE; nonapeptide; cancer; melanoma; breast cancer; HLA;
KW   histocompatability; human leucocyte antigen; probe; treatment;
KW   therapy; vaccine.
OS   Synthetic.
PN   WO9405304-A.
PD   17-MAR-1994.
PF   30-AUG-1993; U08157.
PR   31-AUG-1992; US-938334.
PR   26-MAR-1993; US-037230.
PR   07-JUN-1993; US-073103.
PA   (LUDW-) LUDWIG INST CANCER RES.
PI   Boon-falleur T,  De Plaen E,  Lurquin C,  Traversari C;
PI   Van Derbruggen P;
DR   WPI; 94-100844/12.
DR   N-PSDB; Q44751.
PT   New nona:peptide derived from tumour rejection antigen precursor
PT   - presented by HLA-A1 cancer cells, for use in diagnosis or
PT   therapy of esp. melanoma and breast cancer.
PS   Disclosure; Page 19; 33pp; English.
CC   An isolated nonapeptide having the amino acid sequence Glu-Val-Asp-
CC   Pro-Ile-Gly-His-Leu-Tyr is derived from the tumour rejection antigen
CC   precursor encoded by the MAGE-3 gene and presented by HLA-A1.  The
CC   nonapeptide can be used in a vaccine to treat a cancerous condition
CC   involving HLA-A1 subtype cancerous cells.  The nucleic acid encoding
CC   the nonapeptide can be used as a probe to identify tumour cells.
CC   This sequence is homologous to the peptide described and is encoded
CC   by the MAGE-1 gene.
SQ   Sequence   9 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT    9
ID   R83932 standard; peptide; 9 AA.
AC   R83932;
DT   05-JUN-1996  (first entry)
DE   MHC class I restricted antigenic peptide #2.
KW   MHC class I; antigen; MAGE; melanoma; breast cancer; bladder cancer;
KW   Titermax; cytotoxic T-lymphocyte; tumour; pathogenic disease; bacteria;
KW   parasite; human; animal.
OS   Synthetic.
PN   WO9528958-A1.
PD   02-NOV-1995.
PF   21-APR-1995; U04975.
PR   22-APR-1994; US-233496.
PA   (SLOK ) SLOAN KETTERING INST CANCER RES.
PI   Dyall R,  Nikolic-Zugic J;
DR   WPI; 95-382848/49.
PT   Cytotoxic T-cell induction by MHC class I-restricted peptide in
PT   adjuvant - useful for treating tumours and baeterial or parasitic
PT   pathogenic diseases
PS   Claim 11; Page 38; 50pp; English.
CC   The sequences given in R83931-49 are MHC class I restricted 8-12
CC   amino acid antigenic peptides.  This peptide is derived from MAGE
CC   and is present in melanoma, breast and bladder cancer.  These
CC   peptides may be administered to a subject in combination with a
CC   suitable adjuvant, pref. Titermax (RTM), to induce cytotoxic T-
CC   lymphocytes.  This method may be used in the treatment of a tumour
CC   or a pathogenic disease, esp. diseases of bacterial or parasitic
CC   origin, in humans and animals, e.g monkeys, dogs cows, horses, etc.
SQ   Sequence   9 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   10
ID   R29769 standard; Peptide; 9 AA.
AC   R29769;
DT   22-APR-1993  (first entry)
DE   Antigen E peptide.
KW   Antigen; tumorigenic cell; A+ B+; T-cell; response; syngeneic;
KW   animal; mouse; tumour rejection antigen precusor; TRAP; P1A.
OS   Homo sapiens.
PN   WO9220356-A.
PD   26-NOV-1992.
PF   22-MAY-1992; U04354.
PR   23-MAY-1991; US-705702.
PR   09-JUL-1991; US-728838.
PR   23-SEP-1991; US-764364.
PR   12-DEC-1991; US-807043.
PA   (LUDW-) LUDWIG INST CANCER RES.
PI   Boon T, Chomez P, De Plaen E, Lurquin C, Traversari C;
PI   Van Den Eynde B, Van Der Bruggen P, Van Pel A;
DR   WPI; 92-415460/50.
PT   Nucleic acid mol. encoding a human tumour rejection antigen
PT   precursor - useful as an immunostimulant in a vaccine for
PT   treating and preventing cancers, also useful in diagnosis
PS   Disclosure; Page 97; 142pp; English.
CC   This sequence represents the sequence of the antigen E.  Antigens suc
CC   as this one cause a T-cell response to be elicited which transplanted
CC   into a syngeneic animal, usually a mouse.  This antigen is derived fr
CC   thecell line MEL3.1.  See also Q32351.
SQ   Sequence   9 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   11
ID   W68371 standard; peptide; 9 AA.
AC   W68371;
DT   14-OCT-1998  (first entry)
DE   Human MAGE-1 peptide binds HLA-A1.
KW   Antigen; major histocompatibility complex; MHC; lymphocyte; detection
KW   immobilisation; cytotoxic T-cell; tumour; leukaemia; lymphoma;
KW   viral infection.
OS   Synthetic.
OS   Homo sapiens.
PN   WO9744667-A2.
PD   27-NOV-1997.
PF   21-MAY-1997; F00892.
PR   21-MAY-1996; US-651925.
PA   (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
PA   (INSP ) INST PASTEUR.
PI   Abastado J, Kourilsky P, Langlade-Demoyen P, Lone Y;
DR   WPI; 98-018653/02.
```

PT Detection, purification and elimination of antigen-specific
PT lymphocytes - for producing cytotoxic T cells for immuno-therapy of
PT cancers and viral infection
PS Disclosure; Page 30; 222pp; French.
CC Peptides W68301-W68384 are examples of antigens (Ag) which can be loaded
CC onto recombinantly produced major histocompatibility complex (MHC)
CC molecules in a method of detecting antigen-specific lymphocytes. The
CC MHC-antigen complex is then immobilised on a solid support and a sample
CC containing cells recognising the MHC-Ag complex may be isolated. This
CC peptide is derived from the human MAGE-1 protein and binds the human
CC leukocyte antigen A1 (HLA-A1). A similar method is used to isolate,
CC purify or eliminate Ag-specific T-cells or to produce Ag-specific
CC cytotoxic T-cells (CTC). The method is also used to detect and quantify
CC tumour-specific T-cells and to generate CTC for specific killing of
CC tumour cells (solid tumours, leukaemia or lymphoma) by injection into
CC a human or animal, but also for treating viral infections.
SQ Sequence 9 AA;

Query Match 100.0%; Score 61; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.32e-01;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 EADPTGHSY 9
||||||||||
Qy 1 EADPTGHSY 9

RESULT 12
ID W77125 standard; peptide; 9 AA.
AC W77125;
DT 16-NOV-1998 (first entry)
DE gp75/TRP-1 synthetic peptide epitope 1.
KW Tyrosinase; tyrosinase cytotoxic lymphocyte response;
KW cytotoxic T lymphocyte; cysteine-depleted; melanoma.
OS Synthetic.
PN WO9833810-A2.
PD 06-AUG-1998.
PF 29-JAN-1998; U01592.
PR 30-JAN-1997; US-037781.
PA (UYVI-) UNIV VIRGINIA PATENT FOUND.
PI Engelhard VH, Hunt DF, Kittlesen D, Slingluff CL;
DR WPI; 98-437388/37.
PT Disease specific immunogen - comprises disease specific cytotoxic T
PT lymphocyte epitope used to elicit melanoma specific CTL response
PS Disclosure; Page 27; 93pp; English.
CC The peptide epitope W77119-W77138 were created for human tumour-specific
CC cytotoxic T lymphocyte response. These peptides are are cysteine-
CC depleted mutants of a native disease-specific CTL epitope. The cysteine-
CC depleted CTL epitopes elicit a stronger or more specific CTL response
CC than the native epitope. The epitopes can be used in a disease-specific
CC immunogen to protect a mammal against disease in particular melanomas.
CC The peptides may also be used to screen a sample for the presence of
CC an antigen with the same epitope, or with a different cross-reactive
CC epitope.
SQ Sequence 9 AA;

Query Match 100.0%; Score 61; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.32e-01;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 EADPTGHSY 9
||||||||||
Qy 1 EADPTGHSY 9

RESULT 13
ID W75734 standard; peptide; 9 AA.
AC W75734;
DT 19-NOV-1998 (first entry)
DE Peptidase-resistant peptide 2.
KW Tumour antigen MZ2-E; T-cell; immunotherapy; cytolytic T-cell; CTL;
KW therapeutic agent; peptidase; MZ2-E antigen peptide analogue; HLA;
KW human leucocyte antigen; MHC; lysis; vaccine.
OS Synthetic.
FH Key Location/Qualifiers
FT Misc_difference 2
FT /note= "D-form residue"
FT Misc_difference 8
FT /note= "D-form residue"
PN WO9833511-A1.
PD 06-AUG-1998.
PF 19-NOV-1997; U21296.
PR 05-FEB-1997; US-795733.
PA (CNRS) CENT NAT RECH SCI.
PA (LUDW-) LUDWIG INST CANCER RES.
PI Ayyoub M, Gairin JE, Mazarguil H, Monsarrat B, Van Den Eynde B;
DR WPI; 98-437166/37.
PT Peptidase-resistant peptide(s) that bind to HLA molecules and
PT related antibodies - particularly for treatment of cancer by
PT inducing proliferation of cytotoxic T cells
PS Claim 20; Page 20; 32pp; English.
CC Sequences W75733-W75736 are peptidase-resistant peptides which are
CC analogues of the tumour antigen MZ2-E. This antigen is a potential
CC target for T-cell based immunotherapy and can also be used to stimula
CC the antigen-specific CTL, however its use as a therapeutic agent is
CC limited due to its degradation by peptidase. The MZ2-E antigen pepti
CC analogues were modified at both peptidase sensitive portions, and wer
CC all shown to exhibit a longer half-life relative to peptidase degrada ion
CC as well as the ability to bind a human leucocyte antigen (HLA). The
CC specific peptides W75733 and W75735 were established to have a compar le
CC affinity for the MHC as the tumour antigen, and W75735 was found to b
CC the ideal peptide analog to use due to it also being able to sensitis
CC the target cells to lysis by effector molecules at similar concentrat ns
CC to those of the antigen MZ2-E. These peptide analogues can be used i
CC vaccines to induce an immune response for treating conditions in whic
CC abnormal HLA/peptide complexes are present on the surface of cells.
SQ Sequence 9 AA;

Query Match 100.0%; Score 61; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.32e-01;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 EADPTGHSY 9
||||||||||
Qy 1 EADPTGHSY 9

RESULT 14
ID W00897 standard; Peptide; 9 AA.
AC W00897;
DT 23-MAY-1997 (first entry)
DE Human melanoma MAGE1 tumour associated antigen p161-169.
KW Adeno-associated virus; vector; liposome; transfection;
KW dendritic cell; melanoma; MAGE1; adoptive immunotherapy;
KW tumour associated antigen.
OS Homo sapiens.
PN WO9703703-A1.
PD 06-FEB-1997.
PF 19-JUL-1996; U12012.
PR 21-JUL-1995; US-001312.
PR 01-NOV-1995; US-007184.
PR 01-DEC-1995; US-566286.
PA (RHON) RHONE POULENC RORER PHARM INC.
PI Lebkowski JS, Philip R;
DR WPI; 97-145208/13.
PT Adeno-associated virus:liposome complexes for transfecting dendritic
PT cells - for inducing immune response, useful for treating e.g.
PT neoplasia or infections
PS Example 5; Page 58; 134pp; English.
CC Tumour associated antigens (W13660-61, W00878-903) can be loaded
CC into dendritic cells and used to induce antitumour immunity.
CC Alternatively, the dendritic cells are transfected with adeno
CC associated virus plasmid DNA (which includes DNA encoding the
CC tumour associated antigen) complexed with cationic liposomes. The
CC amtigen loaded or transfected dendritic cells can be used to
CC generate tumour antigen-specific cytotoxic T lymphocytes for use in

```
CC   adoptive immunotherapy in a patient having the corresponding
CC   tumour.  A suitable antigen comprises amino acids 161-169 (W00897)
CC   of human melanoma MAGE1.
SQ   Sequence   9 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   15
ID   W75736 standard; peptide; 9 AA.
AC   W75736;
DT   19-NOV-1998  (first entry)
DE   Peptidase-resistant peptide 4.
KW   Tumour antigen MZ2-E; T-cell; immunotherapy; cytolytic T-cell; CTL;
KW   therapeutic agent; peptidase; MZ2-E antigen peptide analogue; HLA;
KW   human leucocyte antigen; MHC; lysis; vaccine.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   Modified_site   2
FT                   /note= "N-Methyl-Alanine"
FT   Modified_site   8
FT                   /note= "N-Methyl-Serine"
PN   WO9833511-A1.
PD   06-AUG-1998.
PF   19-NOV-1997; U21296.
PR   05-FEB-1997; US-795733.
PA   (CNRS ) CENT NAT RECH SCI.
PA   (LUDW-) LUDWIG INST CANCER RES.
PI   Ayyoub M, Gairin JE, Mazarguil H, Monsarrat B, Van Den Eynde B;
DR   WPI; 98-437166/37.
PT   Peptidase-resistant peptide(s) that bind to HLA molecules and
PT   related antibodies - particularly for treatment of cancer by
PT   inducing proliferation of cytotoxic T cells
PS   Claim 20; Page 20; 32pp; English.
CC   Sequences W75733-W75736 are peptidase-resistant peptides which are
CC   analogues of the tumour antigen MZ2-E.  This antigen is a potential
CC   target for T-cell based immunotherapy and can also be used to stimulate
CC   the antigen-specific CTL, however its use as a therapeutic agent is
CC   limited due to its degradation by peptidase.  The MZ2-E antigen peptide
CC   analogues were modified at both peptidase sensitive portions, and were
CC   all shown to exhibit a longer half-life relative to peptidase degradation
CC   as well as the ability to bind a human leucocyte antigen (HLA).  The
CC   specific peptides W75733 and W75735 were established to have a comparable
CC   affinity for the MHC as the tumour antigen, and W75735 was found to be
CC   the ideal peptide analog to use due to it also being able to sensitise
CC   the target cells to lysis by effector molecules at similar concentrations
CC   to those of the antigen MZ2-E.  These peptide analogues can be used in
CC   vaccines to induce an immune response for treating conditions in which
CC   abnormal HLA/peptide complexes are present on the surface of cells.
SQ   Sequence   9 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   16
ID   R82988 standard; Peptide; 9 AA.
AC   R82988;
DT   26-FEB-1996  (first entry)
DE   P815 antigenic peptide.
KW   P815 antigen; P1A antigen; cancer; vaccine.
OS   Synthetic.
PN   WO9523874-A1.
PD   08-SEP-1995.
PF   23-FEB-1995; U02203.
PR   01-MAR-1994; US-204727.
PR   10-MAR-1994; US-209172.
PR   01-SEP-1994; US-299849.
PR   30-NOV-1994; US-346774.
PA   (LUDW-) LUDWIG INST CANCER RES.
PI   Boon-Falleur T,  Brasseur F,  Chomez P,  De Plaen E;
PI   De Smet C,  Gaugler B, Lethe B,  Marchand M,  Patard J;
PI   Szikora J,  Van Den Eynde B,  Van Derbruggen P,  Weynants P;
DR   WPI; 95-320586/41.
PT   Determn. of cancerous condition(s) - using a nucleic acid as a
PT   primer to determine expression of a MAGE tumour rejection antigen
PT   precursor
PS   Example 13; Page 22; 121pp; English.
CC   Using the sequence of the P815A antigen precursor gene P1A
CC   (T01176), an antigenic peptide (R82988) which was A+B+ (i.e.
CC   characteristic of cells which express both A and B antigens) was
CC   produced.  The peptide lysed PO.HTR cells in the presence of
CC   cytolytic T lymphocyte cell lines, and may be useful as a vaccine
CC   component.
SQ   Sequence   9 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   17
ID   R90692 standard; peptide; 9 AA.
AC   R90692;
DT   31-JUL-1996  (first entry)
DE   Human leukocyte antigen (HLA-A1) presented peptide MZ2-E.
KW   Human leukocyte antigen; HLA-A1; MAGE-1 derived;
KW   blood mononuclear cell; BMC; CD8-beta+ cell; cytolytic T cell;
KW   CTL cell; treatment; tumour cell; diagnosis; assay;
KW   presented peptide.
OS   Synthetic.
PN   WO9535500-A1.
PD   28-DEC-1995.
PF   14-JUN-1995; U07559.
PR   17-JUN-1994; US-261541.
PA   (LUDW-) LUDWIG INST CANCER RES.
PI   Boon-Falleur T,  Coulie P,  Van Der Bruggen P;
DR   WPI; 96-058510/06.
PT   Prodn. of specific cytolytic T cell sub-populations - by contacting
PT   blood mononuclear cells with specific peptide(s) and a population of
PT   CD8-beta(+) cells
PS   Claim 5; Page 19; 25pp; English.
CC   The present peptide is the human leukocyte antigen (HLA-A1), MAGE-1
CC   derived presented peptide, MZ2-E. By contacting a sample of blood
CC   mononuclear cells (BMC) with the peptide (which binds directly to
CC   HLA-A1 mols. on the surface of the BMC) and CD8-beta+ cells (which
CC   stimulate peptide/HLA-A1 complex specific CD8-beta+ cells), a
CC   peptide/HLA-A1 complex specific cytolytic T (CTL) cell
CC   subpopulation can be obtd. . The CTL cells obtd. can be
CC   administered to a patient to treat tumour cell related conditions,
CC   and can be used in diagnostic methods, e.g. in assays for the
CC   peptide/HLA-A1 complex.
SQ   Sequence   9 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9
```

```
RESULT   18
ID   W56729 standard; peptide; 9 AA.
AC   W56729;
DT   31-JUL-1998  (first entry)
DE   MAGE-1 antigenic partial peptide sequence (residues 161-169).
KW   MAGE; replication defective; adenovirus; tumour; antigen; cancer;
KW   immunotherapy; tumour rejection antigen precursor; TRAP; CTL;
KW   human leukocyte antigen; HLA; cytolytic T lymphocyte.
OS   Synthetic.
PN   WO9815638-A2.
PD   16-APR-1998.
PF   06-OCT-1997; U17948.
PR   06-OCT-1996; US-027891.
PA   (LUDW-) LUDWIG INST CANCER RES.
PI   Cerrottini J, Jongeneel CV, Reed DS, Rimoldi D,
PI   Romero P;
DR   WPI; 98-240824/21.
PT   New replication-defective adenoviruses - comprise insert encoding
PT   tumour rejection antigen precursor(s), useful for, e.g. cancer
PT   immuno-therapy
PS   Examples; Page 42; 56pp; English.
CC   This is a partial sequence of the MAGE-1 antigenic peptide used in the
CC   methods of the invention. The specification provides a new nucleic acid
CC   molecule comprising a replication-defective adenovirus genome containing
CC   an insert encoding a tumour rejection antigen precursor (TRAP). The
CC   replication-defective adenovirus genome is useful as a vector for
CC   introducing a TRAP molecule into mammalian (especially human) cells. The
CC   recombinant adenovirus is preferably targetted to tumour cells, e.g. by
CC   binding a ligand to the virus coat. The TRA peptides which are generated
CC   from the expressed TRAP are presented by human leukocyte antigen (HLA)
CC   molecules and as a result cytolytic T lymphocyte (CTL) production is
CC   increased (claimed). The CTL's then kill the TRAP-expressing tumour
CC   cells. Also, cells transfected by the recombinant adenovirus can be used
CC   for assessing the processing of TRAPs, including post-translational
CC   modifications. The adenovirus (genome) can be administered by injection,
CC   topical application or intracavitarily in 106-1010 pfu doses. The range
CC   of TRA peptides produced by replication-defective adenovirus means that
CC   patients with a range of HLA phenotypes can be treated. Also, host cell
CC   immune response to TRA's is enhanced, e.g. by induction of tumour-
CC   specific cytolytic T lymphocytes.
SQ   Sequence   9 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db       1 EADPTGHSY 9
           |||||||||
Qy       1 EADPTGHSY 9


RESULT   19
ID   R75954 standard; Peptide; 9 AA.
AC   R75954;
DT   06-MAR-1996  (first entry)
DE   Melanoma antigen (MAGE-1) epitope.
KW   MAGE-3; melanoma antigen; vaccine; immune response; immunogenic peptide;
KW   cytotoxic T lymphocyte response; CTL; melanoma; breast cancer; antibody.
OS   Homo sapiens.
PN   WO9519783-A1.
PD   27-JUL-1995.
PF   25-JAN-1995; U01000.
PR   25-JAN-1994; US-186266.
PA   (CYTE-) CYTEL CORP.
PI   Celis E,  Grey HM,  Kubo RT,  Sette A;
DR   WPI; 95-269270/35.
PT   Immunogenic peptide(s) that induce immune response to cancer cells
PT   - that express a MAGE-3 protein peptide epitope used in vaccines or
PT   adoptive immuno:therapy to induce cytotoxic T lymphocytes
PS   Example; Page 33; 44pp; English.
CC   R75954 is derived from MAGE-1 protein.  It was used to show the
CC   specificity of CTL response to MAGE-3 peptides shown in R75942-53.
CC   R75942 is derived from the sequence of the melanoma antigen (MAGE-3)
CC   protein and can be used to elicit a primary cytotoxic T lymphocyte
CC   response against cells expressing MAGE-3. Synthetic peptides R75945-5
CC   can be used therapeutically to elicit CTL responses to melanoma, brea t,
CC   colon, prostate, or other cells which express proteins with this epit pe.
CC   The peptides have specific HLA-A1 binding capacity.
SQ   Sequence   9 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db       1 EADPTGHSY 9
           |||||||||
Qy       1 EADPTGHSY 9


RESULT   20
ID   R99343 standard; Protein; 9 AA.
AC   R99343;
DT   22-APR-1997  (first entry)
DE   MAGE-1 nonapeptide.
KW   HLA binding peptide; cell lysis; cytolytic T cell; MAGE family; human
KW   tumour rejection antigen precursor; TRA; MAGE-1; tumour; cancer cell;
KW   antibody; melanoma; universal effector cell; vaccine; breast cancer;  TL;
KW   therapy.
OS   Homo sapiens.
PN   WO9626214-A1.
PD   29-AUG-1996.
PF   01-FEB-1996; U01489.
PR   23-FEB-1995; US-393273.
PA   (LUDW-) LUDWIG INST CANCER RES.
PI   Boon-Falleur T,  De Plaen E,  Gaugler B,  Lurquin C;
PI   Romero P,  Traversari C, Van Den Eynde B,  Van Der Bruggen P;
DR   WPI; 96-402317/40.
DR   N-PSDB; T35408.
PT   New nona:peptide(s) that bind to HLA molecule(s) and induce lysis  -
PT   by specific cytolytic T cells, for diagnosis and treatment of
PT   tumours and to expand T cells in vitro.
PS   Example 4; Fig 4; 41pp; English.
CC   R99343-R99350 represent MAGE nonapeptides, based on the tumour rejecti n
CC   antigen region of the full length MAGE sequences.  These peptides wer
CC   used to design the nonapeptides of the invention (see R99337-R99342),
CC   which bind to a HLA molecule on a cell, and provoke lysis by cytolyti  T
CC   cells (CTLs) specific for a complex of the HLA molecule and nonapepti e.
CC   The nonapeptides can be used diagnostically to identify tumours
CC   expressing a particular HLA molecule, or to identify cancer cells.  T e
CC   peptides can also be used therapeutically, to induce a CTL response t
CC   tumours (where the peptides are optionally coupled to tumour-specific
CC   antibodies), or to induce a response by CTLs that are otherwise inact e.
CC   The peptide sequences may also be used to expand specific CTLs in vit o
CC   for later return to the patient, such as for treating melanoma.  Tumo r
CC   cells can be identified by using DNA encoding the nonapeptides as pro es.
CC   Non-human cells transformed with the HLA-A1 gene and a DNA sequence
CC   encoding one of the peptides, can be used to generate CTLs, or to det t
CC   the presence of CTLs in human samples.  The non-human transformed cel s,
CC   when polytransformed, are universal effector cells, and can be used i
CC   vaccines, or for treating melanoma or breast cancer.
SQ   Sequence   9 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db       1 EADPTGHSY 9
           |||||||||
Qy       1 EADPTGHSY 9


RESULT   21
ID   W78838 standard; peptide; 9 AA.
AC   W78838;
```

DT 17-NOV-1998 (first entry)
DE MAGE-1 protein fragment 161-169.
KW Microparticle; delivery; polymeric matrix; autoantigen; tumour antigen;
KW class II associated peptide; pathogen; gene therapy; genetic disease;
KW infection; downregulation; immune response.
OS Homo sapiens.
OS Synthetic.
PN WO9831398-A1.
PD 23-JUL-1998.
PF 22-JAN-1998; U01499.
PR 06-JAN-1998; US-003253.
PR 22-JAN-1997; US-787547.
PA (PANG-) PANGAEA PHARM INC.
PI Curley JM, Hedley ML, Langer RS, Lunsford LB;
DR WPI; 98-427556/36.
PT New preparations of microparticles - comprising a synthetic polymer
PT matrix and nucleic acid comprising an expression vector for use in
PT gene therapy
PS Disclosure; Page 10; 101pp; English.
CC A microparticle preparation (MP) has been developed, consisting of
CC microparticles having a diameter of less than 100 mu m. The MP comprises:
CC (a) a polymeric matrix (PM) consisting of one or more synthetic polymers
CC having a solubility in water of less that 1 mg/l; and (b) an expression
CC vector selected from RNA molecules (at least 50% of which are closed
CC circles) or circular plasmid DNA (at least 50% of which are supercoiled).
CC Also described is a MP of at most 20 microns in diameter, comprising: (a)
CC a PM; and (b) a NAM comprising an expression control sequence operatively
CC linked to a coding sequence, where the coding sequence encodes an
CC expression product selected from: (1) a polypeptide at least 7 amino
CC acids in length, having a sequence identical to the sequence of: (i) a
CC fragment of a naturally-occurring mammalian protein; or (ii) a fragment
CC of a naturally-occurring protein from an infectious agent which infects
CC a mammal; (2) a peptide having a length and sequence which permits it to
CC bind to an MHC class I or II molecule; and (3) the polypeptide or the
CC peptide linked to a trafficking sequence. W69763 to W69765, and W78793
CC to W78897 are peptide fragments for use in the present invention. The
CC MPs are highly effective vehicles for the delivery of polynucleotides
CC into phagocytic cells. They can be used for gene therapy, e.g. for
CC treating genetic diseases, infections or tumours or for downregulating
CC an immune response.
SQ Sequence 9 AA;

Query Match 100.0%; Score 61; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.32e-01;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 EADPTGHSY 9
|||||||||
Qy 1 EADPTGHSY 9

RESULT 22
ID R65112 standard; peptide; 9 AA.
AC R65112;
DT 06-OCT-1995 (first entry)
DE MAGE 1 immunogenic peptide 161-169.
KW MAGE 1; immunogenic peptide 161-169; cytotoxic C cells;
KW in vitro activation; cancer; AIDS; bacterial infections; malaria;
KW fungal infections; tuberculosis; hepatitis.
OS Homo sapiens.
PN WO9504817-A.
PD 16-FEB-1995.
PF 01-AUG-1994; U08672.
PR 06-AUG-1993; US-103401.
PA (CYTE-) CYTEL CORP.
PI Celis E, Kubo R, Serra H, Tsai V, Wentworth P;
DR WPI; 95-090895/12.
PT In vitro activation of cytotoxic T cells for selected killing of
PT target cells - for treating e.g. cancer, AIDS, hepatitis etc.by
PT incubating them with antigen presenting cells loaded with
PT appropriate immunogenic peptide
PS Example 3; Page 35; 53pp; English.
CC R65109-R65145 are immunogenic peptides, they are used in a new
CC method for the in vitro activation of cytotoxic T cells (CTC).
CC This is achieved by incubating the CTCs with antigen presenting
CC cells loaded with an appropriate immunogenic peptide (e.g. one
CC of the above peptides). By selecting the peptides used the
CC following diseases and infections can be treated; cancer, AIDS,
CC hepatitis, other viral and bacterial infections, malaria and
CC tuberculosis.
SQ Sequence 9 AA;

Query Match 100.0%; Score 61; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.32e-01;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 EADPTGHSY 9
|||||||||
Qy 1 EADPTGHSY 9

RESULT 23
ID R63675 standard; Protein; 9 AA.
AC R63675;
DT 22-JUN-1995 (first entry)
DE Synthetic peptide derived from exon 3.1 of MAGE 1.
KW Melanoma antigen-1; MAGE-1; cytolytic T cells; antigen E; exon 3.1.
OS Synthetic.
PN WO9423031-A.
PD 13-OCT-1994.
PF 17-MAR-1994; U02877.
PR 26-MAR-1993; US-037230.
PA (LUDW-) LUDWIG INST CANCER RES.
PI Boon-falleur T, Gaugler B, Van DEN EYNDE B, Van DER BRUGGEN P;
DR WPI; 94-333192/41.
PT New tumour rejection antigen precursor MAGE3 - useful in
PT treatment and diagnosis of cancer
PS Example 34; Page 36; 105pp; English.
CC R63675 is a synthetic peptide derived from exon 3.1 of melanoma
CC antigen-1 (MAGE-1), it was used to transfer antigen-E cytolytic T
CC lymphocyte sensitivity to normally non-sensitive cells.
SQ Sequence 9 AA;

Query Match 100.0%; Score 61; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.32e-01;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 EADPTGHSY 9
|||||||||
Qy 1 EADPTGHSY 9

RESULT 24
ID R78824 standard; peptide; 9 AA.
AC R78824;
DT 26-MAR-1996 (first entry)
DE MAGE-1 cytotoxic T lymphocyte epitope.
KW MAGE-1; cytotoxic T; CTL; epitope; helper T; HTL; lymphocyte;
KW cell; viruses; parasites; tumours; antigens; disease prevention;
KW treatment.
OS Homo sapiens.
PN WO9522317-A1.
PD 24-AUG-1995.
PF 16-FEB-1995; U02121.
PR 16-FEB-1994; US-197484.
PA (CYTE-) CYTEL CORP.
PI Celis E, Chesnut RW, Grey H, Sette AD, Vitiello MA;
DR WPI; 95-302545/39.
PT Compsn. inducing cytotoxic T lymphocyte response to pref. viral,
PT bacterial, parasitic or tumour antigens - useful in the treatment
PT and prevention of diseases associated with the antigen e.g.
PT hepatitis B
PS Disclosure; Page 17; 109pp; English.
CC A compsn. which induces a cytotoxic T lymphocyte (CTL) response to
CC an antigen (Ag) in a mammal comprises, a CTL Ag response inducing
CC peptide (i.e. R78824-R78853) and a lipid conjugated helper T cell

```
CC   inducing peptide. The compsn. induces a CTL response to bacterial,
CC   viral or tumour Ags, and is therefore useful in the treatment and
CC   prevention of diseases associated with the Ag.
SQ   Sequence   9 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   25
ID   R65135 standard; peptide; 9 AA.
AC   R65135;
DT   09-OCT-1995  (first entry)
DE   MAGE 1 immunogenic peptide A01.
KW   MAGE 1; immunogenic peptide A01; cytotoxic C cells;
KW   in vitro activation; cancer; AIDS; bacterial infections; malaria;
KW   fungal infections; tuberculosis; hepatitis.
OS   Homo sapiens.
PN   WO9504817-A.
PD   16-FEB-1995.
PF   01-AUG-1994; U08672.
PR   06-AUG-1993; US-103401.
PA   (CYTE-) CYTEL CORP.
PI   Celis E,  Kubo R,  Serra H,  Tsai V,  Wentworth P;
DR   WPI; 95-090895/12.
PT   In vitro activation of cytotoxic T cells for selected killing of
PT   target cells  - for treating e.g. cancer, AIDS, hepatitis etc.by
PT   incubating them with antigen presenting cells loaded with
PT   appropriate immunogenic peptide
PS   Example 3; Page 38; 53pp; English.
CC   R65109-R65145 are immunogenic peptides, they are used in a new
CC   method for the in vitro activation of cytotoxic T cells (CTC).
CC   This is achieved by incubating the CTCs with antigen presenting
CC   cells loaded with an appropriate immunogenic peptide (e.g. one
CC   of the above peptides). By selecting the peptides used the
CC   following diseases and infections can be treated; cancer, AIDS,
CC   hepatitis, other viral and bacterial infections, malaria and
CC   tuberculosis.
SQ   Sequence   9 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   26
ID   R49224 standard; Protein; 9 AA.
AC   R49224;
DT   31-AUG-1994  (first entry)
DE   HLA-A1 MAGE 1 antigen peptide fragment 958.01.
KW   Immunogenic; HLA-A3.2; HLA-A1; HLA-A11; binding motif; MHC molecule;
KW   immune response; viral infection; cancer; prostate cancer; lymphoma;
KW   hepatitis; AIDS; antibody; diagnosis; melanoma antigen.
OS   Synthetic.
PN   WO9403205-A.
PD   17-FEB-1994.
PF   06-AUG-1993; U07421.
PR   07-AUG-1992; US-926666.
PR   05-MAR-1993; US-027746.
PA   (CYTE-) CYTEL CORP.
PI   Celis E,  Grey HM,  Kubo RT,  Sette A;
DR   WPI; 94-065403/08.
PT   Peptide which specifically binds selected MHC allele - used to
PT   induce an immune response for treatment or prevention of viral
PT   infection or cancer, or for diagnosis
PS   Eaxmple 16; Page 116; 150pp; English.
CC   The sequences given in R47304-33 and R49201-44 are immunogenic
CC   peptides which have a HLA-A3.2, HLA-A1 or a HLA-A11 binding motif.
CC   These peptides may be used in the composition of the invention.
CC   These peptides are capable of binding selected MHC molecules and
CC   inducing an immune response.  They can be used to treat and/or
CC   prevent viral infection and cancer, eg. prostate cancer, lymphoma,
CC   hepatitis or AIDS.  They can also be used to produce antibodies for
CC   use as diagnostic or therapeutic agents.  The peptides can also be
CC   used as diagnostic agents.
SQ   Sequence   9 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


RESULT   27
ID   R47330 standard; Protein; 9 AA.
AC   R47330;
DT   31-AUG-1994  (first entry)
DE   HLA-A1 MAGE 1 antigen peptide fragment 161-169.
KW   Immunogenic; HLA-A3.2; HLA-A1; HLA-A11; binding motif; MHC molecule;
KW   immune response; viral infection; cancer; prostate cancer; lymphoma;
KW   hepatitis; AIDS; antibody; diagnosis; melanoma antigen.
OS   Synthetic.
PN   WO9403205-A.
PD   17-FEB-1994.
PF   06-AUG-1993; U07421.
PR   07-AUG-1992; US-926666.
PR   05-MAR-1993; US-027746.
PA   (CYTE-) CYTEL CORP.
PI   Celis E,  Grey HM,  Kubo RT,  Sette A;
DR   WPI; 94-065403/08.
PT   Peptide which specifically binds selected MHC allele - used to
PT   induce an immune response for treatment or prevention of viral
PT   infection or cancer, or for diagnosis
PS   Eaxmple 8; Page 52; 150pp; English.
CC   The sequences given in R47304-33 and R49201-44 are immunogenic
CC   peptides which have a HLA-A3.2, HLA-A1 or a HLA-A11 binding motif.
CC   These peptides may be used in the composition of the invention.
CC   These peptides are capable of binding selected MHC molecules and
CC   inducing an immune response.  They can be used to treat and/or
CC   prevent viral infection and cancer, eg. prostate cancer, lymphoma,
CC   hepatitis or AIDS.  They can also be used to produce antibodies for
CC   use as diagnostic or therapeutic agents.  The peptides can also be
CC   used as diagnostic agents.
SQ   Sequence   9 AA;

  Query Match          100.0%;  Score 61;  DB 1;  Length 9;
  Best Local Similarity 100.0%;  Pred. No. 2.32e-01;
  Matches    9;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 EADPTGHSY 9
          |||||||||
Qy      1 EADPTGHSY 9


Search completed: Wed Sep 13 06:33:38 2000
Job time : 40 secs.
```